The Genetic Aetiology of Ectopic Maxillary Canine Teeth

Camilleri, Simon

Awarding institution:
King's College London

The Genetic Aetiology of Ectopic Maxillary Canine Teeth

PhD

King's College London

Simon Camilleri M.Sc. (Lond.), L.D.S.R.C.S.(Eng.), M. Orth. R.C.S.(Eng.)

F.D.S.R.C.S (Eng.)

December 2012

24,921 words

Acknowledgements

My thanks go to my supervisors Fraser McDonald and Chris Scerri and the staff at the, Molecular Genetics Laboratory at the University of Malta. To those whose patience, help and advice seemed endless: Kathryn Lewis for the segregation analysis, Franz Ruschendorf for the linkage analysis, Bert Eussen for the CNV analysis, Michael Simpson for the Exome Scan, Ozahama Atsushi for the *in situ* hybridisation and Paul Sharpe, whose support made the latter two investigations possible. To Joseph Borg for all his advice and encouragement. I am especially grateful to my wife Naila, who had to put up with me throughout.

Chapter 1 Introduction and Literature Review 15
1.1 Introduction 15
1.1.1 Aetiology of ectopic canines 15
1.1.2 Evidence for a genetic aetiology 16
1.1.3 Environmental influence 18
1.1.4 Local prevalence 18
1.1.5 Founder effect 19
1.1.6 Diagnosis of Ectopic Canines 20
1.1.7 Sequelae of untreated ectopic eruption 21
1.1.8 Associated dental anomalies 22
1.2 Mechanisms of tooth eruption 22
1.2.1 Eruption of the maxillary canine 22
1.2.2 Intraosseous movement of the tooth 23
1.2.3 Control of tooth eruption by dental follicle 24
1.2.4 Osteoblast regulation 24
1.2.5 Osteoclast regulation 25
1.2.6 Osteoblast osteoclast interaction 28
1.2.7 Genetic disorders and tooth eruption 30
1.3 Identification of the disease gene 33
1.3.1 Familial aggregation 33
1.3.2 Segregation Analysis 33
1.3.3 Identification of the genetic component 34
1.3.4 Positional Cloning 34
1.3.5 Genetic maps 35
1.3.6 Two point vs multipoint analysis 36
1.3.7 LOD scores 36
1.3.8 Parametric linkage analysis 37
1.3.9 Locus heterogeneity 38
1.3.10 Linkage disequilibrium 39
1.3.11 Allele sharing methods 40
1.3.12 Allele sharing linkage analysis 40
1.3.13 Homozygosity mapping 41
1.3.14 Association analysis 41
1.3.14.1 Case-Control Methods 42
1.3.14.2 Family-based Methods 42

1.3.15 Combined linkage and association analysis 44
1.3.16 Fine mapping 44
1.3.17 Copy Number Variations and Loss Of Heterozygosity 45
1.3.18 Genetic markers 46
1.3.19 Next Generation Sequencing 48
1.4 Disease segregation 49
1.4.1 Heritable dental anomalies and EC 50
1.5 Study Design 53
1.5.1 Segregation Analysis 53
1.5.2 Multipoint mapping 53
1.5.3 Choice of marker 54
1.5.4 Whole Exome Sequencing 55
1.5.5 In situ hybridisation 55
Chapter 2 Aims of the Study 56
2.1 Aims and objectives 56
2.2 Null Hypothesis 56
2.3 Potential Benefits 57
Chapter 3 Materials and Methods 58
3.1 Segregation analysis 58
3.1.1 Data analysis 59
3.2 Linkage analysis 60
3.2.1 Genotyping 67
3.2.2 Quality Control 68
3.2.3 Compilation of pedigrees 68
3.2.4 Data analysis 69
3.2.5 Genotyping Quality Control 70
3.2.6 Linkage Disequilibrium 70
3.2.7 Haplotyping 71
3.2.8 Exclusion of linkage 71
3.2.9 Copy Number Variations 71
3.3 Whole Exome Sequencing 72
3.3.1 Exome capture and library preparation 72
3.3.2 Massive parallel sequencing 73
3.3.3 Stepwise filtering approach for exome sequencing 74
3.3.4 Sanger Sequencing 75

3.4 In situ hybridisation 76
Chapter 4 Results 84
4.1 Segregation analysis 84
4.1.1 Discussion 89
4.1.1.1 Dental Anomalies 89
4.1.1.2 Pedigree Analysis 89
4.1.1.3 Segregation Analysis 91
4.1.1.4 Twin study 91
4.1.2 Conclusions 93
4.2 Genome wide linkage analysis 93
4.2.1 Results 93
4.2.1.1 LD quality control 94
4.2.1.2 Analysis of V8 94
4.2.1.3 Analysis of V7 95
4.2.1.4 Haplotyping 96
4.2.1.5 Exclusion of linkage 96
4.2.1.6 Copy Number Variations 101
4.2.2 Discussion 101
4.2.2.1 LD and quality control 101
4.2.2.2 Genetic model 102
4.2.2.3 Phenotype 104
4.2.2.4 Results of haplotyping 105
4.2.2.5 Identification of a candidate gene 105
4.2.2.6 Exclusion of linkage 106
4.2.2.7 Copy number variations 106
4.2.3 Conclusions 107
4.3 Exome Sequencing 107
4.3.1 Exome capture 108
4.3.2 Limitations of Exome Sequencing 109
4.3.3 Filtering 110
4.3.4 Results 112
4.3.4.1 Sanger Sequencing results. 113
4.3.5 Discussion 117
4.3.5.1 Quality control 118
4.3.5.2 Filtering 118

4.3.5.3 Candidate genes 120
4.4 In situ hybridisation 124
4.4.1 Results 125
4.4.2 Discussion 126
4.5 Conclusion 130
4.5.1 Future investigations 131
References 133
Appendix 1. Ethical Approval 163
Appendix 2. Tables and Pedigrees 171
Appendix 3. Publications 188

Tables

Table 1. List of genetic conditions where delay or failure of eruption is a feature. The full list may be found as Table 22, in Appendix 2. 31

Table 2. LOD scores for multipoint linkage analysis as suggested by Lander and Krugylak (1995). 37

Table 3. Specifications of the Affymetrix SNP 6.0 chip (http://www.affymetrix.com/browse/products.jsp?productId=131533&navMode=34000&navAction=jump&aId=productsNav#1_1). 48

Table 4. The developmental dental anomalies associated with EC. 50

Table 5. The effects of AXIN2, PAX9 and MSX1 mutations on the human dentition. 51

Table 6. Primers for Investigation of Candidate Genes. 76

Table 7. Conditions for PCR. The same program was used for all four primer pairs. 76

Table 8. The proportions of ectopic canine-affected subjects used in calculating the relative risks. 85

Table 9. The -2 log likelihoods of the various models tested. The wild-type allele frequency and the penetrance of the genotypes AA, AB and BB were computed for the various models. Chi square tests were used to compare models with a difference of no more than one degree of freedom. 88

Table 10. The pairwise concordance rates of the monozygotic and dizygotic twins/triplets in the sample. 88

Table 11. Causes for monozygotic twin discordance (Silva et al 2011). 92

Table 12. The change in size and position of the peak with a spaced marker set 94

Table 13. Positive LOD scores Pedigree V8. 95

Table 14. Positive LOD scores Pedigree V7 95

Table 15. Linkage exclusion of eruption related genes. 96

Table 16. Linkage exclusion of SNPs flanking those identified by Geller et al. 101

Table 17. Summary statistics for exome sequencing, giving the variants in each individual as well as those common to both. 114

Table 18. Nonsynonymous novel variants common to the two affected members. The LOD scores of both V7 and V8 analyses indicate the areas are unlikely to be shared IBD. 115

Table 19. The rare variants that fell within areas identified by linkage analysis. The variant on Chromosome 20 has a positive LOD score in the conservative V8 analysis but has a negative score in the V7 analysis. 115

Table 20. The final list of candidate genes. A SIFT score of <0.05 predicts the mutation to be deleterious. PolyPhen scores range from 0 (benign) to 1 (deleterious). 116

Table 21. The SNVs in PPP1R14C exon 1 in eighteen unrelated EC cases (P1-P18). One SNV is novel and rs2297672 has a MAF of 0.39. N = variant absent, Y = variant present. . 117

Table 22. Genetic conditions where failure of eruption is a feature. 171

Table 23. The variants with a frequency of 1% or less in the 1000 genomes database which were common to both exomes. 174

Figures

Figure 1. Radiograph showing ectopic maxillary and mandibular canines and retained deciduous teeth with no permanent successors. 17

Figure 2. Population Growth of the Maltese Islands. The Islands were constantly ravaged by North African corsairs after the first millennium and up to the early 1500's the population never grew above the 10,000 mark. The arrival of the Knights of Malta brought an influx of about 10,000 knights, soldiers and dependents. The stability offered by a military base allowed the population to grow rapidly and now exceeds 400,000, with a relatively minor foreign input. Data taken from Blouet (2004). 20

Figure 3. The pathways to a candidate gene. 35

Figure 4. The family tree used in the study. The pedigree numbers are outside the symbol, the age of the subject is inside. Males are square symbols, females round. Individual #41 was categorised as a phenocopy and so is not marked as affected. 62

Figure 5. The phenotypes of the affected members of the family. EC was present in all members. IPH was present in all affected members except #23 and #41. The canine in #41 was buccal and this individual was categorised as a phenocopy. 63

Figure 6. Relative risk dropoff as a function of degree of relation. In diseases where the parameter λR -1 decreases by a factor of 2 for each decreasing degree of unilineal relationship, this is suggestive of a single locus. (λR = Relative Risk). 86

Figure 7. (A) V8 LOD plot. The maximum LOD from this pedigree was 1.5 and maximum scores were seen on chromosomes 4, 5, 6, 11, 13 and 20. (B) V7 LOD plot. The maximum LOD from this family was 3.9 and scores of 2.6 and 2.5 were seen on chromosomes 11 and 20 respectively. 97

Figure 8. The haplotype on Chromosome 6. The disease haplotype is in red, inherited from the maternal founder. Markers have been selected for illustration, not all are shown. 98

Figure 9. The haplotype on Chromosome 11. The disease haplotype is in red, inherited from the maternal founder. Markers have been selected for illustration, not all are shown. 99

Figure 10. The haplotype on Chromosome 20. The disease haplotype is in red, inherited from the paternal founder. Markers have been selected for illustration, not all are shown.....100

Figure 11. The distribution of the variants in the family. The SNVs in *ANO5* and *PPP1R14C* together are necessary to produce the phenotype. 123

Figure 12. *ANO5* in situ. There is weak or no expression in dental tissue in the initial stages of tooth formation, however there is diffuse, weak expression in both tooth and dental follicle in the later stages of tooth formation. The yellow dots outline the developing tooth germ indicated by the arrows. Scale 200 µM.....127

Figure 13. *PPP1R14C* in situ. There is strong expression in dental tissues throughout tooth formation. There is particularly strong expression in ameloblasts and weak expression in the follicle and supporting tissues at the bell stage of formation. The yellow dots outline the developing tooth germ indicated by the arrows. Scale 200µM. 128

List of Abbreviations

ANO5 Anoctamin 5

Arsb Arylsulphatase B

ATP Adenosine Triphosphate

Bcor BCL6 Co-Repressor

CCD Cleidocranial Dysplasia

Clcn7 Chloride Channel 7

cM Centimorgan

CNV Copy Number Variation

Col1α1 Collagen Type1 alpha 1

Col1α2 Collagen Type1 alpha 2

COST European Cooperation in Science and Technology

CSF1 Colony Stimulating Factor 1

CTP Cytidine Triphosphate

DEPC Diethylpyrocarbonate

Dhof Focal Dermal Hypoplasia

DNA Deoxyribonucleic Acid

DPT Dental Panoramic Tomogram

DTT Dithiothreitol

EC Ectopic Canines

Ed1 Ectodysplasin 1

Edar Ectodysplasin 1 Anhydrotic Receptor

Edaradd EDAR Associated Death Domain

EDTA Ethylenediaminetetraacetic acid

EGF Epidermal Growth Factor

EGFR Epidermal Growth Factor Receptor

EPM2A EPM2A gene, Laforin

Ercc6 Excision Repair Cross Complementing 6

ERK Extracellular Signal-regulated Kinase

EtOH Ethyl Alcohol

Evc Ellis Van Creveld

FGF3 Fibroblast Growth Factor 3

Fgfr1 Fibroblast Growth Factor Receptor 1

Fgfr2 Fibroblast Growth Factor Receptor 2

Flna Filamin A

Gnas1 Guanine Nucleotide Binding Protein

GDP General Dental Practitioner

GTP Guanosine Triphosphate

HDAC Histone Deacetylase

Ids Iduronate 2 Sulphatase Deficiency

IL1A Interleukin 1-A

IPH Incisor-premolar hypodontia

LD Linkage disequilibrium

Lmna Lamin A

LOH Loss Of Heterozygosity

MAF Minor Allele Frequency

MAPD Median of the Absolute values of all Pairwise Differences

MEK Mitogen Activated Protein Kinase Kinase

MCP1 Monocyte Chemotactic Protein

MERLIN Multipoint Engine for Rapid Likelihood INference

MSX1 Muscle Segment Homeobox 1

MSX2 Muscle Segment Homeobox 2

Nemo NF Kappa B Essential Modulator

Nipbl Nipped B-Like

Oscs Osteopathia Striata with Cranial Sclerosis

PBS Phosphate Buffered Saline

PCI Phenol-Chloroform-Isoamyl alcohol

PCR Polymerase Chain Reaction

PPP1CA Protein phosphatase 1, catalytic subunit, alpha isoform

PPP1R14C Protein Phosphatase1, Regulatory Subunit14C

ProK Proteinase K

r^2 Coefficient of Linkage Disequilibrium

RANBP9 RAN-Binding Protein 9

RANKL Receptor Activator of Nuclear Factor Kappa B Ligand

RANKL Receptor Activator of Nuclear Factor Kappa B

RCT Randomised Controlled Study

RFLP Restriction-fragment length polymorphism

RNA Ribonucleic Acid

RT Reverse Transcription

Runx2 Runt-related Transcription Factor 2

SDS Sodium dodecyl sulphate

SNP Single Nucleotide Polymorphism

SOX10 SRY Box 10

SSC Saline-Sodium Citrate

STR Short Tandem Repeat

SYN3 Synapsin III

TAE buffer Tris-Acetate-EDTA buffer

Tcirg1 T Cell Immune Regulator 1

TE Buffer Tris-EDTA buffer

TEA Triethanolamine

TEN Buffer Tris-EDTA-NaCl buffer

TNFα Tumor Necrosis Factor Alpha

Tris Tris(hydroxymethyl)aminomethane

UTP Uridine Triphosphate

VNTR Variable Number Tandem Repeat

Whsc1 Wolf-Hirschhorn Syndrome Candidate 1

Gene, mRNA and protein conventions.

The gene nomenclature follows the HGNC and MGI guidelines in that human genes symbols are italicised, with all letters in uppercase. Protein symbols are not italicised and are uppercase. Mouse genes are italicised with the first letter capitalised and the protein product not italicised and in uppercase. Where it is necessary to distinguish, the symbols are preceded with (gDNA) (cDNA) and (mRNA) for genomic DNA. copy DNA and mRNA respectively.

Abstract

Introduction. The ectopic canine (EC) is a common clinical complication of dental development appearing in 1-2% of the Western population The aetiology is controversial with opinion divided as to a genetic or environmental mechanism. This study addresses the hypothesis that genetic factors play an important role in the aetiology of ectopic maxillary canines. Elucidation of the extent of genetic factors will determine the feasibility of further molecular studies to identify putative genes responsible for ectopic eruption and aid in their identification. Molecular control of tooth eruption would reduce or eliminate the need for surgical procedures associated with buried and impacted teeth and facilitate treatment of those dentofacial anomalies where failure of tooth eruption is a feature.

Methods. The study is divided into five parts: 1. A segregation analysis was carried out on 63 pedigrees where a proband was identified as affected with EC, in order to determine whether a genetic component does exist and to provide parameters for further investigation by linkage analysis. 2. Following a positive result from the segregation analysis, linkage analysis was carried out on DNA obtained from an informative, three generation family with seven affected members. 3. Whole exome sequencing was carried out on two distantly related affected members of this family, common, novel and rare variants being identified. 4. The exons and intron-exon junctions of the candidate genes were sequenced using Sanger sequencing in the family and in 18 unrelated cases of EC. 5. In situ hybridisation was carried out using the genes *ANO5* and *PPP1R14C*.

Results. The segregation analysis identified a major genetic component with autosomal dominant transmission and the likelihood of a single major locus being involved. The linkage analysis identified several regions of interest and this data was used to filter the results of the exome sequencing. The presence of variations in both PPP1R14C and ANO5 were necessary to precipitate the phenotype. Sanger sequencing of unaffected family members and of unrelated cases showed no similar variants. In situ hybridisation showed both PPP1R14C and ANO5 to be expressed in tooth and supporting tissues, leading to a supposition of digenic inheritance.

Conclusion. The genes PPP1R14C and ANO5 are implicated in the aetiology of EC in a digenic inheritance pattern in this family. Further sequencing of affected families and functional studies are required as well as investigation of the methylation status of discordant monozygotic twins.

Chapter 1 Introduction and Literature Review

1.1 Introduction

The ectopic canine (OMIM 189490) is a clinical complication of dental development. (Figure 1). It appears in 1-2% of the Western population (Ericson and Kurol, 1986, Grover and Lorton, 1985, Shah et al., 1978, Thilander and Jakobsson, 1968) and in 4-7 % of the Maltese population (Camilleri, 1995, Camilleri et al., 1995). The condition has been reported to be more common in females than in males in roughly a 2:1 ratio. Ectopic maxillary canines have been found in a skull dating from 7250 B.C (Iseri and Uzel, 1993). Nelson reported a high prevalence of ectopic maxillary canines in an isolated prehistoric population (Nelson, 1992). Colyer described the condition in a medieval English population, with 5 specimens out of 166 (3%) exhibiting palatally displaced canines (Colyer, 1922).

1.1.1 Aetiology of ectopic canines

The aetiology of ectopic canines is probably heterogenic and most likely multifactorial. A number of causes have been put forward over the years, such as the long path of eruption, narrow or large arches, crowding, lack of guidance by the lateral incisor root, cystic enlargement of the follicle and familial tendency (Coulter and Richardson, 1997, Bass, 1967, Zilberman et al., 1990, Jacoby, 1983, Thilander and Jakobsson, 1968). A genetic basis has been suggested (Svinhufvud et al., 1988, Bjerklin et al., 1992, Peck et al., 1994). The racial variation, female preponderance, familial occurrence, and association

with other dental anomalies has been attributed to a polygenic aetiology (Kotsomitis and Freer, 1997).

A controversy exists over the aetiology of ectopic canines. Peck and Peck support a genetic theory (Peck et al., 1994) while Becker et al. propose that this anomaly arises from environmental factors, namely lack of mechanical guidance by the lateral incisor root (Becker et al., 1981).

1.1.2 Evidence for a genetic aetiology

The preponderance of evidence supports the view of Peck and Peck (1994).Other dental anomalies are seen to occur with ectopic canines. There is a clear association between this phenomenon, tooth agenesis and tooth malformations, particularly involving the lateral incisors (Moss, 1972, Becker et al., 1984, Svinhufvud et al., 1988, Oliver et al., 1989, Brin et al., 1993, Pirinen et al., 1996). Svinhufvud et al. (1988) demonstrated an association of canine malposition to hypodontia[1] (Svinhufvud et al., 1988). Pirinen et al. (1996) showed that the aetiology of ectopic canines is genetic and is related to hypodontia (Pirinen et al., 1996). Camilleri et al. carried out a complex segregation analysis on a series of families with ectopic canines and found that the data indicated a major genetic component with a single causative locus. The best fitting model was that of a single dominant gene with autosomal dominant inheritance (Camilleri et al., 2008). The reported range of bilateral occurrence of ectopic canines varies between 17 and 45% (Ericson and Kurol,

[1] The term hypodontia is taken to mean agenesis of up to six teeth.

1988, McKay, 1978, Nordenram and Stromberg, 1966, Power and Short, 1993).

Figure 1. Radiograph showing ectopic maxillary and mandibular canines and retained deciduous teeth with no permanent successors.

The prevalence rates of bilateralism for other dental anomalies under genetic control, such as missing maxillary lateral incisors(Davis, 1987)and maxillary canine-first premolar transposition(Peck et al., 1993) are similar, being 20-46% and 23-43 % respectively. The gender ratio shows a bias towards females (Becker et al., 1981, Ericson and Kurol, 1988, McKay, 1978, Nordenram and Stromberg, 1966, Power and Short, 1993, Zilberman et al., 1990). The ratios range from M1:F1.3 to M1:F3.2. These ratios compare favourably with other dental anomalies of genetic origin, the gender ratio of hypodontia being quoted as M1:F1.3 to M1:F1.6 (Davis, 1987) and that for maxillary canine-first premolar transposition M1:F3.8 (Peck et al., 1993).

1.1.3 Environmental influence

There is however, evidence of an environmental component. Timely extraction of the deciduous canines will ameliorate the condition (Ericson and Kurol, 1988, Power and Short, 1993). There is support for environmental factors such as lack of guidance by the lateral incisor root although the developmental absence of lateral incisors often allows the canine teeth to erupt into the line of the arch (Becker et al., 1981, Chaushu et al., 2002, Chaushu et al., 2003b). Excess space in the dental arch may play a part (Paschos et al., 2005). Both genetic and environmental factors may be involved (Ely et al., 2006). Camilleri et al. presented a series of monozygotic twins with ectopic canines. The concordance of these twins was low and equivalent to that of a similar series of dizygous twins. Furthermore, 85% of the three-generation families in this study showed instances where an obligate carrier exhibited a normal phenotype, although the condition was transmitted. This is consistent with an environmental or an epigenetic component (Camilleri et al., 2008).

1.1.4 Local prevalence

This condition is a clinical problem on the Island of Malta having an impact on orthodontic services. Patients with ectopic maxillary canines constitute a major load of specialist practice treatments (eighteen per cent of a private workload, of which 14% required fixed appliance treatment in combination with oral surgery [data not shown]). The problem is similarly evident at the School Dental Clinic in Mater Dei Hospital. A random survey of 270 DPTs taken of individuals between the ages of eight and eighteen years of age found 19% to

have EC. The prevalence of hypodontia in this sample was 29% (unpublished data). In a sample of 168 patients with ectopic canines from a private practice, the incidence of hypodontia is just over 20%. These figures are well above those quoted for orthodontic populations (Rose, 1966, Sacerdoti and Baccetti, 2004). The incidence of lateral incisor hypodontia in the general population is also higher than average (Camilleri and Mulligan, 2007) and is significantly different to the prevalence of lateral incisor hypodontia in a European population (Bot and Salmon, 1977).

1.1.5 Founder effect

The prevalence of the associated conditions of ectopic canines and hypodontia appears to be higher in Malta than elsewhere. If a genetic aetiology is present, one explanation of this phenomenon may be the ‘founder effect’. The Maltese population has grown dramatically over the past 500 years, from 17,000 in 1535 to over 400,000 in the twenty first century. (Blouet, 2004) Any particular genes carried by the founders may well be over represented in the present population (Figure 2). In small populations the relative importance of genetic drift is higher; deleterious alleles can become more frequent and 'fixed' in a population due to chance. This makes the Maltese population a good subject for investigation of EC as the chances of genetic heterogeneity are lower and linkage disequilibrium surrounding disease genes may be higher.

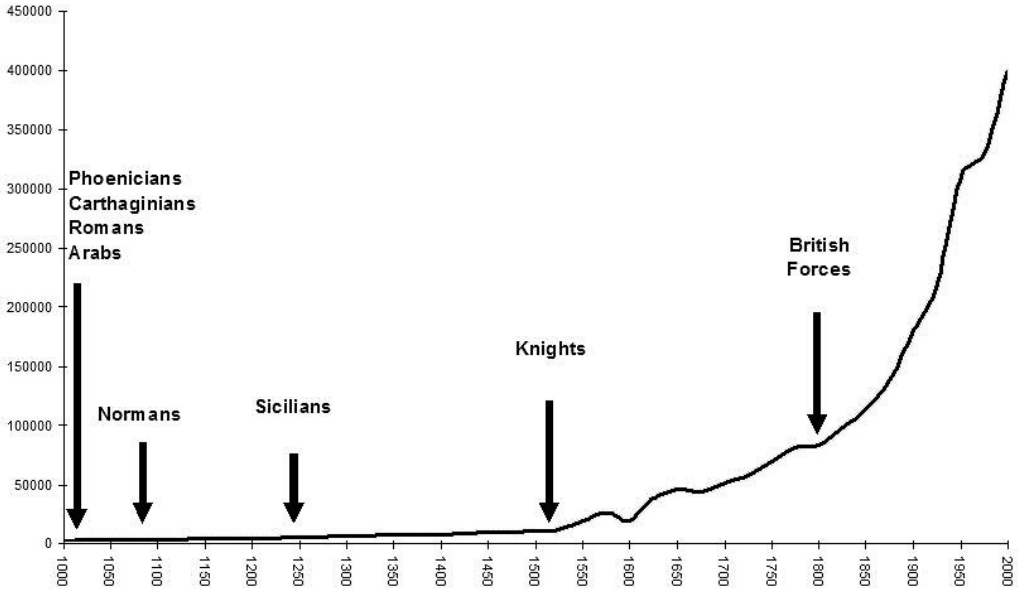


Figure 2. Population Growth of the Maltese Islands. The Islands were constantly ravaged by North African corsairs after the first millennium and up to the early 1500's the population never grew above the 10,000 mark. The arrival of the Knights of Malta brought an influx of about 10,000 knights, soldiers and dependents. The stability offered by a military base allowed the population to grow rapidly and now exceeds 400,000, with a relatively minor foreign input. Data taken from Blouet (2004).

1.1.6 Diagnosis of Ectopic Canines

The permanent canines are usually palpable by age 9 or 10. Inability to palpate the canines at this age is an indication for radiographic localisation of the permanent canines. However this method is not entirely satisfactory as palpation of the canines requires a degree of clinical experience. Many general dental practitioners (GDP) may not possess the skills to evaluate the clinical presentation. Furthermore, as these teeth are among the last to erupt at 12-13 years, the tendency is to await the eruption and in doing so, miss the opportunity for timely referral to a specialist. At present, the accepted interceptive treatment is limited to extraction of the deciduous canines in selected cases. This encourages the permanent canines to erupt into their

correct place. By this time though, the permanent canines are already erupting along an ectopic path. There is no scientific evidence to support this method. No randomised controlled studies (RCT) of adequate quality have been published to date (Parkin et al., 2009). The success rate varies, with 62-78% of cases being reported as normalised (Ericson and Kurol, 1988, Power and Short, 1993). Only mild cases stand a good chance of correction. The earlier this procedure is carried out, the better the prognosis. However, as the condition is painless and not obvious to the layman, a large proportion of cases are identified either by chance or by loss of a deciduous tooth in the adult years, leading the individual to seek treatment very late on in the development of the dentition. Early identification is essential if these patients are to be treated successfully and fixed appliance treatment is necessary to correct established cases.

1.1.7 Sequelae of untreated ectopic eruption

Failure to diagnose and treat EC may lead to palatal or buccal placement of the tooth, leading to difficult orthodontics or the loss of the canine. This would require complex prosthetic replacement. Other complications include loss of space for the tooth by drift of adjacent teeth, further complicating any orthodontics or prosthetic replacement. Dentigerous cyst formation may result in 1.44% of impacted teeth (Mourshed, 1964) and the ectopic canine may cause root resorption of the adjacent central and lateral incisor teeth, this occurring in 38% of laterals and in 9% of centrals (Ericson and Kurol, 2000)

1.1.8 Associated dental anomalies

Garn et al demonstrated the existence of a genetic interrelationship among tooth agenesis, systematic tooth size reduction and generalized retardation of tooth development, three of the abnormalities that appear associated with ectopic canines (Garn et al., 1963, Garn et al., 1964, Garn and Lewis, 1970). Both Bjerklin and Baccetti found an increased prevalence of other developmental anomalies associated with ectopic canines, notably tooth agenesis, tooth size reduction, ankylosed deciduous molars, other ectopically positioned teeth and supernumerary teeth (Bjerklin et al., 1992, Baccetti, 1993, Baccetti, 1998). Shalish et al. demonstrated an association between ectopic canines, ankylosed deciduous molars and ectopically erupting mandibular premolars (Shalish et al., 2009, Shalish et al., 2010)

It is unlikely that EC is an isolated phenomenon, evidence points to it being a genetic anomaly frequently occurring in association with other genetically interrelated dental anomalies.

1.2 Mechanisms of tooth eruption

1.2.1 Eruption of the maxillary canine

The maxillary permanent canine starts to calcify at 1½ years of age, between the roots of the first primary molar. As the jaws grow, the canine moves apically, away from the first primary molar. The first premolar then develops in the same site as the canine (Broadbent, 1941). As the jaws grow rapidly in depth and width, the teeth move to maintain their correct relation to each other. As the maxilla grows, the first premolar moves distally relative to the canine,

providing space for the canine to erupt. This involves precise co-ordination of movement of the tooth germs in the growing maxilla. This movement is most likely effected by osteoblast-osteoclast interaction, controlled by the dental follicle, as part of the eruption process. A deficiency in the cell signalling process of one tooth or more adjacent teeth, at an early stage, could well cause the tooth buds to move in the wrong direction (or fail to move), leading to ectopic eruption or impaction of the tooth.

1.2.2 Intraosseous movement of the tooth

The intraosseous stage of tooth eruption involves bone resorbtion to form an eruption pathway and interradicular bone formation, root growth and fundic bone apposition which move the erupting tooth into the eruption pathway. The primary determinant of both the direction and rate of tooth eruption is the rate of formation of the eruption pathway and it's co-ordination with bone formation in selected areas of the crypt and alveolar crest (Schroeder et al., 1992, Wise et al., 2007). These events are regulated by the dental follicle, which develops regions to initiate and control bone formation, regions to initiate and control bone resorbtion and regions which remain neutral. Where these regions develop in a particular follicle determines the direction of tooth eruption and when they develop. It also determines the time of eruption and how they are synchronised and will determine whether there will be complications (Marks and Schroeder, 1996).

1.2.3 Control of tooth eruption by dental follicle

The dental follicle initiates and controls tooth eruption by inducing bone resorbtion occlusal to the tooth and bone deposition at the base of the crypt (Marks and Cahill, 1987, Wise et al., 2007). The coronal half of the follicle produces bone resorbing factors such as CSF-1 and RANKL while the basal half produces BMP2, a bone formation factor (Wise et al., 2005, Wise and Yao, 2006). The tooth germ is carried to the surface by bone deposition at the base of the crypt and bone resorbtion ahead. Communication takes place between tooth germs and bone-forming/resorbing cells, synchronising the two processes, perhaps to ensure correct spatial positioning of teeth in the jaws (Ohazama et al., 2004, Gao et al., 2003). CSF-1 reduces expression of OPG in cultured dental follicle cells (Wise et al., 2000) as well as reducing the amount of OPG secreted (Wise et al., 2004).

1.2.4 Osteoblast regulation

Bone tissue consists of hydroxyapatite crystals and various kinds of extracellular matrix proteins including type I collagen, osteocalcin, osteonectin, osteopontin, bone sialoprotein, and proteoglycans (Mundlos et al., 1997, Young et al., 1992, Robey et al., 1993). Most of these bone matrix proteins, together with hydroxyapatite crystals, are secreted and deposited by polarised mature osteoblasts, which are aligned on the bone surface.

Runx2 null mice (Komori et al., 1997) have no bone tissue, osteoblasts or osteoclasts. *Runx2* dominant negative mice display an osteopenic phenotype

(Ducy et al., 1999). This indicates that RUNX2 is essential for osteoblast commitment from undifferentiated mesenchymal cells.

RUNX2 has been shown to be essential for normal bone formation, with perturbation of bone formation if levels are insufficient. However RUNX2 inhibits osteoblast differentiation at a late stage (Liu et al., 2001).

Adult transgenic mice overexpressing Runx2 showed osteopenia with a decrease in bone mineral density due to enhanced osteoclastogenesis (Geoffroy et al., 2002). This is consistent with evidence that *Runx2* is autoregulated in part by negative feedback on its own promoter to stringently control expression and function during bone formation. (Drissi et al., 2000). Osteoblasts from synostosed sutures exhibited an increase in RUNX2 expression and activity, possibly explaining the enhanced proliferation and bone-forming ability of these cells (Shevde et al., 2001).

RUNX2 expression is affected by a diversity of signalling pathways. Binding of the extracellular matrix (ECM) to cell surface integrins; mechanical loading; FGF2; PTH; and BMPs all influence RUNX2 dependent transcription. These act via the mitogen-activated protein kinase (*MAPK*) and protein kinase A and C (*PKA*, *PKC*) pathways, activating RUNX2 by phosphorylation. This gene plays a central role in coordinating multiple signalling pathways affecting osteoblast differentiation (Franceschi et al., 2003, Franceschi and Xiao, 2003).

1.2.5 Osteoclast regulation

Osteoclastogenesis in the alveolar bone, essential for the accommodation of normal tooth development and eruption, is mediated by RANK-RANKL

signalling (Suzuki et al., 2004). The spatiotemporal pattern and relative abundance of CSF-1, RANKL and its antagonist OPG, during tooth eruption are key determinants of site-specific osteoclast activity in bone surrounding the tooth (Heinrich et al., 2005).

Runx2 affects osteoblast-osteoclast interaction. CSF-1 is secreted from osteoblasts and provides a survival signal to osteoclast precursors and osteoclasts (Tsurukai et al. 2000).

RANKL is expressed in committed preosteoblastic cells and acts by increasing transcription via the RUNX2 binding sites on the *Rankl* promoter, however the ability of osteoprogenitor cells to support osteoclast formation decreases with their maturation stage (Manolagas 2000). Osteoclastogenesis is strongly induced by undifferentiated stromal marrow cells, which produce high RANKL levels. As maturation proceeds, RANKL levels drop and OPG levels rise (Gori et al., 2000).

The action of RUNX2 on the *Rankl* promoter region is affected by Histone Deacetalase (HDAC), which condenses chromatin, making it less accessible to transcription machinery. At the same time, RUNX2 does have a slightly positive effect on the basic promoter activity of the *Rankl* gene (Mori et al. 2006).

Evidence suggests that RUNX2 is a key regulator of RANKL expression at least in chondrocytes. RUNX2 may indirectly affect the pattern of RANKL expression by inducing selective and sequential expression of other signalling and transcription factors which act on the *Rankl* promoter (Kishimoto et al., 2006).

The RANKL antagonist OPG is strongly expressed in *Runx2*$^{-/-}$ calvarial cell lines (Enomoto et al., 2003, Gao et al., 1998). However RUNX2 binding elements are also present in the promoter region of the *Opg* gene. RUNX2 increases the activity of the *Opg* promoter, suggesting that RUNX2 also regulates osteoclastogenesis by inducing the expression of OPG (Thirunavukkarasu et al., 2000).

Forced production of soluble RANKL was found to be insufficient for the complete rescue of osteoclast differentiation in *Runx2*$^{-/-}$ mice, suggesting the presence of another requirement for osteoclast differentiation. Also, treatment of RUNX2-deficient calvarial cells with 1,25$(OH)_2$D3 affected both RANKL and OPG expression and induced osteoclastogenesis (Notoya et al., 2004).

Thus, while RANKL is necessary for osteoclast formation, it is not sufficient. Furthermore the requirement for RUNX2 may be bypassed, at least *in vitro*, with expression of RANKL and OPG and initiation of osteoclastogenesis being influenced via alternative pathways.

Overexpression of RUNX2 will also affect osteoblast-osteoclast interaction. Apert's and Crouzon syndromes are due to a mutation on *FGFR2*, which causes increased affinity for FGF2 .This results in overproduction and overactivity of RUNX2 (Kim et al., 2006, Wilkie and Morriss-Kay, 2001). These related syndromes are also characterized by ectopic and unerupted teeth and delayed dental development (Kaloust et al., 1997, Wilkie and Morriss-Kay, 2001).

Overexpression of RUNX2 increases osteoblast number but inhibits their terminal maturation, resulting in accumulation of less mature osteoblasts

and consequent osteopenia. The increase in immature osteoblast numbers will lead to a dearth of osteoclast-inducing osteoprogenitor cells. (Liu et al., 2001).

1.2.6 Osteoblast osteoclast interaction

The experiments of Cahill and of Marks (Marks and Cahill, 1984, Cahill and Marks, 1980) have conclusively demonstrated that tooth eruption is dependent on osteoclastic activity to form an eruption pathway for the tooth to move inside the bone. Teeth apparently have a ‘window of opportunity’ to erupt. Should this slot be missed, eruption may fail or the tooth may erupt ectopically (Marks et al., 1983). Experiments in rodents where osteoclastic activity is temporarily disrupted produce a picture which is remarkably similar to the clinical condition in humans, namely delayed and ectopic eruption of teeth together with ankylosis and odontome formation (Cielinski et al., 1994, Kodama et al., 1991, Huang and Wise, 2000, Zou et al., 2003, Yoda et al., 2004). Furthermore, in the clinical situation, tooth eruption in humans is often accelerated by localised chronic inflammation. All too often a premolar will erupt, sometimes years ahead of schedule (and other premolar teeth), beneath a chronically infected deciduous molar tooth (O'Meara, 1966).Should a sinus be present, the permanent tooth will often follow the track and erupt buccally (Camm and Schuler, 1990). Osteoclasts are activated by the products of inflammation, with proteins such as TNF-α IL1A, CSF1 and MCP1 being common to both inflammation and tooth eruption. TNF-α and IL1A increase the production of RANKL (Bezerra et al., 2005). Inducing inflammation close to the follicle of the ectopic tooth may alter the direction of tooth eruption.

Extraction of the deciduous tooth will often correct the path of eruption of the ectopic permanent tooth (Ericson and Kurol, 1988, Power and Short, 1993). The closer the crown of the affected tooth is to the root of the deciduous tooth, the more likely the procedure is to succeed. It is possible that the correction of ectopic eruption is due to reactivation of osteoclasts by cytokines produced by surgically induced inflammation and the subsequent local increase in bone turnover as the socket heals.

Premature extraction of a deciduous tooth has a similar effect on the eruption of its successor, even if erupting normally. Eruption is speeded up initially, only to return to normal some time after, once the extraction socket has healed. Indeed, the healed socket will then retard eruption (Loevy, 1989). Exposure of the crown of an ectopic tooth will stimulate eruption of the tooth. The success of this procedure is dependent on the age of the individual (Altonen and Myllarniemi, 1976). Exposure in adult patients often results in a small amount of eruption, possibly due to the transient inflammation produced or the conversion of the follicle to periodontal ligament. That eruption does not proceed further may be due to the lack of active growth of the alveolar process.

Examination of serial radiographs of patients with unilateral ectopic maxillary canines shows that the ectopic eruption pattern may sometimes be evident at or shortly after crown formation. The affected tooth does not upright itself at age 9 (Fernandez et al., 1998, McSherry and Richardson, 1999) but the crown tends to move further mesially than vertically, in contrast with other maxillary canines. The crown will therefore recede from the occlusal plane as alveolar height increases and migrate towards the midline. The apex will migrate

distally, indicating that root formation is proceeding faster than tooth movement (Marks and Schroeder, 1996). This indicates a failure of formation of the eruption pathway of the tooth. The frequent deformation of the apices of these teeth supports this hypothesis. Ankylosis of deciduous teeth is associated with ectopic movement of canines and the prevalence of other ectopic teeth is also increased (Bjerklin et al., 1992, Baccetti, 1998, Baccetti, 2000, Chaushu et al., 2003b). Ankylosis takes place when the cementum of a tooth fuses with the bone of the socket. This could either be due to ectopic ossification of the ligament or due to failure of osteoclasts or dentinoclasts to remove the bony bridge. Cementum formation may also be affected. Histological examination of sections of impacted canines show acellular cementum to be deficient and cellular cementum to be virtually absent (Giuliana et al., 1995).

1.2.7 Genetic disorders and tooth eruption

The weight of evidence points to the causative gene being one which affects bone turnover. Given the complex nature of bone turnover, the possibility of genetic heterogeneity must be considered. A list was compiled of known genetic disorders where ectopic eruption, failure or delay of eruption was a feature (Wise et al., 2002) (Table 1). The common feature running through all these disorders is their disruptive effect on bone remodelling, be it through loss of bone specific transcription factors, bone structural elements, cellular function, or a more fundamental effect on overall growth and development. This is in keeping with the theory of Cahill and Marks that tooth eruption is regulated by coordinated bone apposition and resorbtion. This makes it

Table 1. List of genetic conditions where delay or failure of eruption is a feature. The full list may be found as Table 22, in Appendix 2.

Condition	OMIM	Phenotype	Gene	Mode of Inheritance	Gene function/Pathogenesis
Aperts Syndrome	#101200	Ectopic teeth, delay/failure of eruption	*FGFR2*	Autosomal dominant	Preferentially expressed in osteogenesis
Cleidocranial dysplasia	#119600 Dental variant .0011	Ectopic teeth, delay/failure of eruption, supernumerary teeth	*RUNX2*	Autosomal dominant	Osteoblast-specific transcription factor, regulator of osteoblast differentiation
Primary Failure of eruption	#125350	Localised failure of eruption	*PTHR1*	Autosomal dominant	Activation of phospholipase and inhibition of adenylcyclase through stimulation of inhibitory G proteins
Hutchinson-Gilford Progeria Syndrome	#176670	Delayed and abnormal tooth eruption and morphology	*LMNA*	Autosomal dominant	Control of nuclear architecture and function

unlikely that there is a single gene wholly responsible for the eruption of teeth, as bone turnover is the product of several major genes and many more minor ones. In those syndromes where ectopic teeth are a prime feature, such as Apert's syndrome and Cleidocranial Dysplasia, the genes affected, *FGFR2* and *RUNX2*, are those directly regulating bone turnover. Moreover, *FGFR2* is upstream of *RUNX2* (Choi et al., 2005). Wolf-Hirschorn syndrome, whose dental manifestations are delayed eruption of permanent teeth, retained deciduous teeth and ectopic teeth, is caused by partial deletion of terminal portion of short arm of chromosome 4. Mild forms have been correlated with 4p16.3 deletions. It is worth noting that the homeobox gene *MSX1* on located on 4p16.1, *FGF3* (expressed in the enamel knot, and a downstream gene of *RUNX2*) (Aberg et al., 2004) on 4p16.3, *EVC* on 4p16. FGF's are regulators of bone formation, *MSX1* is a major homeobox gene, upstream of *RUNX2* and *EVC* plays an essential role in skeletal development.

Not all bone remodeling anomalies cause delayed or ectopic eruption. Paget's Disease of bone, a genetically heterogeneous condition caused by mutations which increase production or reduce inhibition of Receptor Activator of Nuclear Factor Kappa B Ligand (RANKL), an osteoclast differentiation factor, does not seem to adversely affect the eruption pathway, the observed clinical effect being premature loss of deciduous teeth in the juvenile variant. Presumably the increase in osteoclastic activity leads to an acceleration, but not disturbance, of the eruption of the permanent dentition. The inference here is that ectopic eruption of teeth is a disease caused by a localised disruption of the eruption pathway, possibly either by failure of the follicle to induce sufficient

osteoclasts to accumulate over the tooth germ or to induce bone deposition apically, or both. Inflammatory products may rescue this failure of eruption. Several gene products may be involved. The condition is not transmitted in a purely Mendelian fashion and the magnitude of the genetic and environmental components is unclear. Thus the condition of ectopic canines may be described as a complex disease.

1.3 Identification of the disease gene

1.3.1 Familial aggregation

In order to establish the genetic component of a disease, a pedigree analysis will determine whether familial aggregation of affected members is evident. Pedigrees are identified on a case-control basis and the proportion of affected members divided by the population prevalence. This gives a figure known as the Relative Risk. The higher this figure, the more indicative it is of a genetic component. However familial aggregation may occur for other reasons than genetic and further investigation is required prior to undertaking molecular studies.

1.3.2 Segregation Analysis

Segregation analysis is the next step, with a large number of pedigrees systematically collected and ascertainment bias catered for. Segregation analysis uses Maximum Likelihood models to fit the pedigree data to various genetic scenarios. Dominant, recessive, polygenic and sporadic (no genetic component) models may all be tested and the most likely model determined.

Penetrance, population allele frequency and heritability, or the proportion of cases of a disease which can be attributed to a genetic effect, may also be calculated.

1.3.3 Identification of the genetic component

A positive result from the segregation analysis is a basis to proceed to molecular studies. There is no specific protocol to determine disease genes, as long as the end result is a short list of plausible genes which can then be examined for mutations. There are four basic methods of establishing this list; i.e. functional cloning, the candidate gene approach, positional cloning and next-generation sequencing (Figure 3). Functional cloning and the candidate gene approach require prior knowledge of either the biochemical basis of the defect or of the function of the gene. These methods are nowadays considered obsolete.

1.3.4 Positional Cloning

Here the position of the gene is determined purely by reference to its physical position on the gene map. The first step in positional cloning is to carry out genome wide linkage analysis or association studies using markers spaced evenly throughout the genome. This may narrow the search down to a specific region which may then be examined for possible candidates. By genotyping genetic markers and studying their segregation through pedigrees, it is possible to infer their relative positions throughout the genome. The discovery of Restriction Fragment Length Polymorphism (RFLP) markers and consequently the more polymorphic, mini and microsatellite DNA markers have permitted

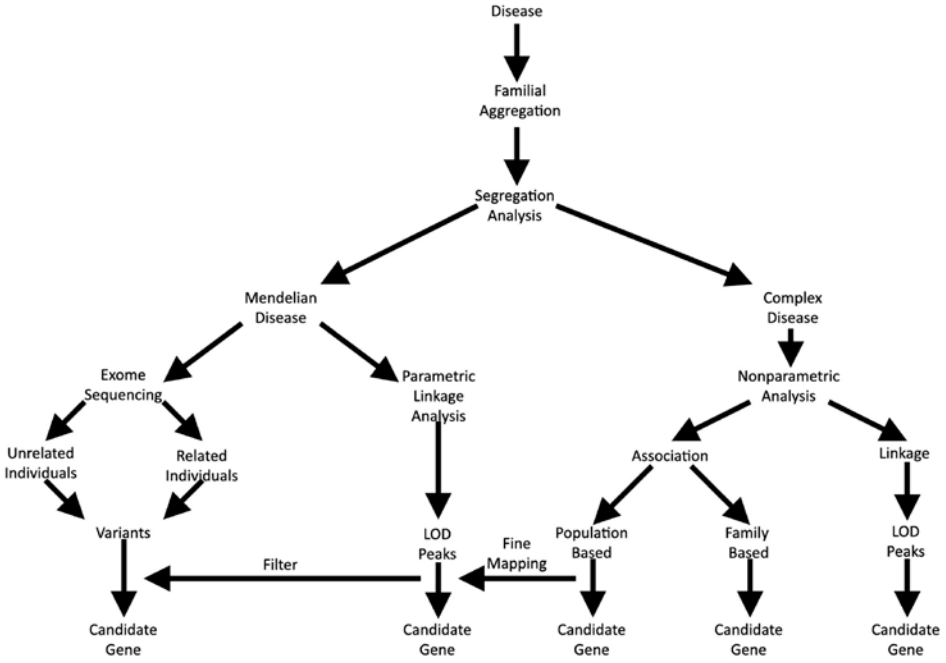


Figure 3. The pathways to a candidate gene.

the identification of several disease genes. The development of high density Single Nucleotide Polymorphism (SNP) arrays allows the use of SNPs to act as highly efficient markers for the analysis of Mendelian diseases (Sellick et al., 2004). Advances in molecular technology have permitted the positional cloning of regions tracked by this technique and have allowed the identification of the mutations responsible. This method works very well for diseases with Mendelian or near-Mendelian inheritance.

1.3.5 Genetic maps

The number of crossovers between two loci is an indication of the distance between them. This is the basis of genetic maps. As recombination frequency is not constant throughout the genome, genetic distances do not correspond

exactly to physical distances but the basic map unit, the Centimorgan, is roughly equal to 1 megabase (Mb) (Dib et al., 1996).

1.3.6 Two point vs multipoint analysis

Two point analysis compares the probability of the data assuming the disease locus is unlinked, to the probability of the data assuming the disease locus is at a specific location, i.e. each locus is analysed separately. Multipoint analysis uses a map of genetic markers to reconstruct inheritance along a chromosome and calculates LOD scores by comparing each possible location to an unlinked locus. The incorporation of a genetic map allows the chromosomal order of linked loci to be established and the maximum inheritance information can be deduced from the available marker data. Multipoint linkage analysis is therefore more efficient than two-point analysis. This is particularly relevant to genotyping with single nucleotide polymorphisms, (SNP) as these are biallelic and consequently less informative than other markers. Multipoint analysis is susceptible to linkage disequilibrium (LD) and this must be taken into account when using dense SNP arrays.

1.3.7 LOD scores

The significance threshold for two point linkage analysis has traditionally been set at a LOD score of 3.0. More stringent criteria are necessary for multipoint analyses in order to reduce the chance of false positive reports and ranges have been suggested for multipoint LOD scores (Table 2) (Lander and Kruglyak, 1995).

Table 2. LOD scores for multipoint linkage analysis as suggested by Lander and Krugylak (1995).

Scale	Description
Suggestive linkage	Statistical evidence that would be expected to occur no more than one time at random in a genome scan. This would include LOD scores from 2.2 to 3.5.
Significant linkage	Statistical evidence that would be expected to occur no more than 0.05 times at random in a genome scan. This would include LOD scores from 3.6 to 5.3
Highly significant linkage	Statistical evidence that would be expected to occur no more than 0.001 times at random in a genome scan. This would include LOD scores over 5.4
Confirmed linkage	Significant linkage that has been confirmed in a further sample, preferably by an independent group of researchers.
Areas of interest	LOD scores below 2.2
Exclusion of linkage	LOD scores below -2

1.3.8 Parametric linkage analysis

Traditional linkage analysis is known as parametric analysis. Parametric methods are more powerful than allele sharing methods, as they examine alleles inherited from a known common ancestor, i.e. are identical by descent (IBD). In other words the marker allele and disease locus are genetically linked. Parametric analysis requires extended pedigrees with multiple affected members for optimal performance. LOD scores may be summed across different families in order to achieve significance. It permits refined recombination mapping in order to localise disease genes and can deal with locus heterogeneity. Parametric methods are most powerful when used to detect single rare alleles with strong effects. Parametric linkage analysis may be used to investigate several patterns of inheritance, however is dependent on a genetic model and requires the model to be well defined. As complex

diseases encompass several often unknown parameters, an incorrect model may affect the estimation of the recombination fraction (Clerget-Darpoux et al., 1986). A number of strategies have been developed to overcome the problem of uncertain inheritance. These involve multiple testing with the necessity to adjust the significance level of the LOD score (Risch, 1991, MacLean et al., 1993, Greenberg et al., 1998). However the exact performance of these strategies in the case of multilocus models is unclear.

Parametric 'affected only' analysis examines the likelihood of inheritance IBD between affected members only. This eliminates the need to specify the penetrance; however the mode of inheritance and allele frequency is still required.

1.3.9 Locus heterogeneity

Locus or genetic heterogeneity is a confounding factor in parametric analysis, as this will affect the LOD score if summed over a number of different families, therefore genetic homogeneity is preferable (Cui et al., 2010). Careful study of the phenotype and selection of families may reduce the problem and several linkage analysis programs incorporate a model which can deal with a limited amount of heterogeneity.

One other way of avoiding this problem is to use a small, relatively isolated population which has recently rapidly expanded from a bottleneck. The small number of founders and the relatively short timeframe results in an increased frequency of founder effects and less mutational heterogeneity.

1.3.10 Linkage disequilibrium

The algorithm used in most multipoint analysis packages assumes independent segregation between markers, i.e. random crossing over. This is not the case as alleles which are physically close together are unlikely to be separated by crossover events. Furthermore, crossing over at one point inhibits similar events in the vicinity and crossing over is not evenly distributed, chromosomes exhibiting both ‘hot’ and ‘cold’ breakage spots. These factors contribute to linkage disequilibrium (LD), which may be defined as the non-random association of alleles at adjacent loci. On a population level, LD is eroded by recombination, therefore older populations, such as Africans; exhibit less LD than other populations (Reich et al., 2001). LD has been shown to be present in isolated, and therefore younger, populations; however the amount of LD exhibited is variable. It is more likely to occur in the case of rare alleles, e.g. founder mutations (Kruglyak, 1999).

Multipoint linkage analysis programs assume linkage equilibrium between genetic markers. (Abecasis et al. 2002). It has been demonstrated that the algorithm can give misleading results in the presence of linkage disequilibrium, especially if founders are not genotyped (Schaid et al. 2002). LD is a recognised problem in multipoint analysis and must be addressed, either by using programs which deal with LD or by preselecting markers not in LD. Linkage disequilibrium however, is the basis of allele sharing methods described below.

1.3.11 Allele sharing methods

Allele sharing methods, incorrectly known as nonparametric or model-free analyses, make no prior assumptions about the disease. These methods examine alleles that are shared between affected individuals and do not take into account unaffected individuals. Here alleles which are identical by state (IBS) are analysed to see if their assortment deviates from that which would be expected to occur were the trait and marker to be in linkage equilibrium. Alleles which are IBS may have the same DNA sequence, but they are not necessarily derived from a known common ancestor. In the case of rare disease alleles in related individuals however, IBS may be taken to imply IBD. Allele sharing analyses are more useful in the case of complex disease where multiple loci each exert small effects along with environmental factors and the mode of inheritance is unclear. These analyses may be linkage or association based.

1.3.12 Allele sharing linkage analysis

Nonparametric linkage analysis requires family data where more than one member is affected, such as affected sibling pairs; the data being assessed to see whether the inheritance of alleles IBD differs from that which would be expected by chance. Extended pedigrees are not a requirement and may raise problems in assessment of data from multiple related individuals (Ott et al., 2011).

1.3.13 Homozygosity mapping

This is a technique that may be used in the investigation of rare recessive disorders. It relies on the detection on stretches of homozygosity common to the affected individuals. There is therefore no need to reconstruct the whole pedigree, or identify consanguineous matings.

1.3.14 Association analysis

Association analyses may be population-based case-control or family-based analysis. Here the frequencies of genetic markers are compared between unrelated affected cases and unaffected controls; i.e. to look for LD between a disease gene and a flanking genetic marker. The case-control method is useful for complex diseases where it would be unlikely to find pedigrees with sufficient affected members.

Association methods make use of recombination events in a population sample. As the shared ancestry of a mutation in a population extends much further back than that of any one pedigree, the increased number of recombination events allow for finer localisation of disease markers. The use of a control sample will further increase resolution, as the ancestry of the unmutated allele extends much further back and comparison of the groups may further localise the disease allele. SNPs are the markers of choice and microarray chips have been developed which carry close to a million SNP markers, with a resolution of 1-200 kilobases (Kb).

The small genetic effects exerted by the target loci means that large sample sizes are required in order to achieve sufficient power to detect association.

The use of large numbers of individuals makes these analyses susceptible to the effects of population stratification, where Type 1 errors (false positives) are the result of different allele frequencies as a consequence of ethnic mixing of the population sample and not due to the disease. Family-based association analysis is immune to this; however collecting sufficient numbers to achieve significance may be impractical (Risch and Merikangas, 1996).

1.3.14.1 Case-Control Methods

The case-control method is the method of choice for analysis of complex diseases where it would be unlikely to find pedigrees with sufficient affected members (Risch and Merikangas, 1996). Furthermore, in the case of late-onset disorders, compilation of pedigrees is not always possible. However, the small genetic effects exerted by the target loci means that large sample sizes are required in order to achieve sufficient power to detect association, leading to Type 1 errors. Finally, independent replication of results, the gold standard of genetic studies, is problematic (Ioannidis et al., 2001). Hurdles in the path of successful replication are the need for an even larger sample size and if carried out in other populations, the difference in both variant frequencies and LD patterns may make the signal undetectable (Clerget-Darpoux and Elston, 2007).

1.3.14.2 Family-based Methods

Family-based association analysis is immune to population stratification as members of a pedigree share a common genetic background. Small pedigrees

such as parent-offspring trios or affected sibling pairs are preferable. Extended pedigrees are not a requirement and may raise problems in assessment of data from multiple related individuals (Ott et al., 2011). Family-based designs also allow for parent of origin effects such as imprinting (Weinberg, 1999). However collecting sufficient numbers to achieve significance may be impractical (Risch and Merikangas, 1996).

Both types of association analyses have their merits, and there are no set criteria for optimal selection of a strategy. Both designs give overall similar estimates of association when compared (Evangelou et al., 2006) and in practice the choice rests mainly on the sampling conditions as well as the mode of inheritance (Cui et al., 2010).

The analysis of data depends on the study design. Chi squared tests, logistic regression or odds ratios may be used in simple cases, however statistical packages are freely available. A P-number is generated which gives the probability of the marker and disease locus being in LD.

Complex disease is much commoner than Mendelian disease. Allele sharing methods are independent of the need to specify a genetic model. Furthermore, other covariates such as environmental input may be factored in. Association analysis reports have largely superseded linkage analysis in the literature over the past decade, their flexibility in these respects contributing to their popularity.

1.3.15 Combined linkage and association analysis

The two techniques may be used to reinforce each other. Nonparametric linkage analysis may be carried out as a replication study to confirm suggestion of association. The transmission disequilibrium test is a family-based test is unique in that it performs a matched analysis of untransmitted versus transmitted alleles thus performing both a test of association and of linkage (Dudbridge, 2007).

1.3.16 Fine mapping

The number of markers required for a genome wide linkage scan is in the order of 400-1000. Linkage analysis peaks are generally in the order of several Mb in width. Further analysis may be carried out either by genotyping more individuals or with the use of a secondary map with more markers, less widely interspersed within the area of interest. Nonetheless, as fine mapping here is dependent on recombination events, the finest resolution is rarely less than 1 Mb (Boehnke, 1994). This may encompass an enormous number of genes, which will require filtering by other means. Candidate genes may be selected for sequencing, as in the discovery of the *PTHR1P* gene's role in Primary Failure of Eruption (Decker et al., 2008), however this may be tedious and expensive. The dense markers used in association analysis will increase the resolution and narrow the interval to be investigated. The precision obtained will depend on the age of the mutation, marker density and sample size.

1.3.17 Copy Number Variations and Loss Of Heterozygosity

A copy number variation (CNV) is a segment of DNA that is 1 kb or larger and is present at a variable copy number in comparison with a reference genome.. CNVs with a population frequency of > 1% are called copy number polymorphisms or CNPs (Feuk et al., 2006). CNVs may be inherited in a Mendelian fashion, similar to SNPs and may often be in strong LD with each other (McCarroll et al., 2008). In theory, dense SNP arrays may detect CNVs by analysing the number of copies of that SNP and arranging them in chromosomal order, however it seems that the data clusters poorly and causes assays to produce Mendelian inconsistencies and violate Hardy-Weinberg equilibrium (McCarroll, 2008). SNP arrays may also be used for virtual karyotyping and will provide information in cases of copy number neutral loss of heterozygosity (LOH) (Mei et al., 2000). Current technology involves the use of dedicated copy number probes hybridised with SNP arrays, though these platforms may still have difficulty in detecting small copy gains or losses. CNVs affect less than 0.5% of the genome but are responsible for a large part of human variation; though evaluating the data from CNV analyses is not straightforward and the data is susceptible to a number of artefacts such as sex ratios and tissue type. A high false discovery rate is common. Furthermore, it seems that common CNVs available on current platforms are unlikely to play a major role in the genetic basis of common diseases (Craddock et al., 2010, Wang et al., 2012). As the disease model is that of Mendelian inheritance, it is possible that CNVs may play a part in the aetiology of EC and linkage analysis results may be used to filter the data.

1.3.18 Genetic markers

A genetic marker should be locus specific, polymorphic and easily genoptyped. There are two types of genetic marker used in linkage analysis nowadays. Short tandem repeats (STRs) are a variant of microsatellite DNA consisting of short, repeated sequences of DNA, 2-5 bases long. The number of repeats varies considerably between individuals and therefore the markers are polymorphic, giving high information content (Evans and Cardon, 2004). SNPs are single-base variations in the DNA sequence. The vast majority of SNPs are biallelic and so are not as informative on an individual basis. There are advantages and disadvantages to using either SNPs or STRs as markers.

On a one to one basis, SNPs are less informative than microsatellites because of their biallelic nature. This may be a disadvantage for association and linkage analysis. Polymorphism information content which is important in linkage analysis is one half to one third lower for SNPs than for STRs.

Closely linked SNP markers with strong linkage disequilibrium (LD) will generate false positives. Many linkage analysis programs based on the Lander-Green algorithm assume linkage equilibrium between genetic markers. (Abecasis et al. 2002). It has been demonstrated that this algorithm can give misleading results in the presence of linkage disequilibrium, especially if founders are not genotyped (Schaid et al. 2002).

Multipoint linkage analysis is extremely sensitive to genotyping error and error rates as small as 1% can significantly decrease the power to detect loci (Abecasis et al. 2001). An increase in marker density also increases the number

of genotyping errors present in the data and the net effect may actually be a decrease in the power to detect linkage. Besides, the biallelic nature of SNPs reduces the ability of software packages to detect genotyping errors.

However SNPs have several advantages over STRs. SNPs are much cheaper to genotype, in terms of cost and labour and they are associated with a lower genotype error (John et al., 2004). A dense map of SNPs provides greater linkage resolution. Thus savings can be made where fine mapping is required.

SNPs are much commoner than STRs and therefore are much more closely placed. This increase in density makes them more informative than conventional microsatellites. LOD peaks which are not detectable by microsatellite scans may be ascertained by using dense SNPs (John et al., 2004). SNPs are simpler to assay and also have a low mutation rate, therefore are more stable over generations.

SNPs therefore do have drawbacks however but are cheaper and more efficient as markers than microsatellites and consequently are in wider use. The Affymetrix Genome-Wide Human SNP Array 6.0 chip (Affymetrix, USA) contains over 906,600 SNPs spaced evenly throughout the genome (Table 3).

Table 3. Specifications of the Affymetrix SNP 6.0 chip (http://www.affymetrix.com/browse/products.jsp?productId=131533&navMode=34000&navAction=jump&aId=productsNav#1_1).

SNPs	906,600
Non Polymorphic Copy Number Probes	946,000
SNP Average Inter-Marker Spacing (bp)	3230
Copy Number Probe Average Inter Marker Spacing (bp)	3160
Combined average coverage (bp)	1600
Average Minor Allele Frequency (Caucasian)	19.6%
Average Heterozygosity (Caucasian)	26.7%

1.3.19 Next Generation Sequencing

Linkage analysis has been successful in identifying many disease genes. However even under ideal circumstances the disease locus can rarely be mapped to an area of less than 1 Centimorgan (cM). An area this size may contain hundreds of genes and conventional Sanger sequencing entails considerable time and expense. The emergence of next-generation massively parallel sequencing and high throughput sequence capture methods has made possible the capture of the whole human genome. However whole-genome sequencing is limited in throughput and too expensive to investigate significant numbers of individuals. It also produces vast amounts of data that present problems in storage and computational power in analysis. The targeted sequencing of all known exons is far more cost effective. Protein coding regions constitute only 1% of the total human genome but are responsible for 85% of mutations (Choi et al., 2009). Mutations having major genetic effects generally disrupt protein coding regions or exon splice sites, mutations in

regulatory regions tend to have lesser effects (Ng et al., 2009) . Thus exome sequencing is a highly efficient method of capturing a large percentage of genomic variation and has reversed the recent decline in interest in monogenic, Mendelian disorders (Antonarakis and Beckmann, 2006). Nonetheless, a large amount of data is generated which must be filtered to narrow the search down to a manageable number of candidate genes. The polymorphisms are filtered first against a control set, obtained either from public databases, assuming that there are no disease alleles present, or from a control cohort of individuals known to be unaffected. This leaves the 'novel' polymorphisms as potential candidates.

Further filtering may be applied by comparing of affected individuals, either on a whole exome basis if unrelated or by comparison of areas IBD in related persons. The remainder of the polymorphisms may be prioritised by prediction of the functional consequences using programs such as SIFT (Kumar et al., 2009) or PolyPhen2 (Adzhubei et al., 2010).

1.4 Disease segregation

One possible method of identifying the cause of EC is to use a trait which segregates out with the disease/trait as a marker for EC. There is a strong clinical association between ectopic canines, tooth agenesis and other dental anomalies, leading to the assumption that the causative loci may be in linkage disequilibrium. This approach has been successful in other genetically isolated populations such as Finland (Jorde, 1995). Case-control association analysis using the responsible polymorphism may lead to the gene(s) responsible for

ectopic eruption and the use of preselected markers would reduce genotyping costs. The strong presence of founder effects in the Maltese population makes this approach more feasible. There are a number of developmental anomalies associated with EC (Table 4).

Table 4. The developmental dental anomalies associated with EC.

Anomaly	Phenotype	Gene responsible	Reference
Incisor-Premolar Hypodontia	Missing lateral incisors and lower second premolars	Unknown	Svinhufvud et al., 1988; Pirinen et al., 1996
Oligodontia	Multiple missing teeth	*MSX1* *PAX9* *AXIN1*	Vastardis et al., 1996; Neubuser et al., 1997; Lammi et al., 2004
Delayed formation and eruption of teeth	Retarded tooth formation, delayed premolar eruption	Unknown	Garn et al., 1961; Rune and Sarnas, 1974; Becker and Chaushu, 2000
Reduction in tooth size and form	Generalised reduced mesiodistal width, small or peg-shaped lateral incisors	Unknown	Grahnen, 1956; Garn and Lewis, 1970; Becker et al., 1981; Baccetti, 1998; Langberg and Peck, 2000; Paschos et al., 2005; Artmann et al., 2010
Ectopic eruption of other teeth	Ectopic first molars, lower second premolars	Unknown	Bjerklin et al., 1992; Shalish et al., 2010
Enamel Hypoplasia	Hypoplastic enamel on central incisors and molars	Unknown	Baccetti, 1998; Bartolo et al., 2010

1.4.1 Heritable dental anomalies and EC

Ideally, the variations recognised as being linked with these heritable dental anomalies would be used to locate and identify the gene responsible for EC as

these would be in linkage disequilibrium and indeed genes responsible for tooth agenesis have been identified in humans. However the experimental evidence is difficult to apply to the clinical situation. The genes known to be associated with non-syndromic tooth agenesis are *MSX1, PAX9* and *AXIN2* (Vastardis et al., 1996, Neubuser et al., 1997, Lammi et al., 2004).

Table 5. The effects of AXIN2, PAX9 and MSX1 mutations on the human dentition.

Gene	Primary Dentition affected	Permanent Incisors affected	Premolars affected	Permanent molars affected	Reference
AXIN2	*	**	***	**	(Lammi et al., 2004)
AXIN2	*	**	***	**	(Bergendal et al., 2011)
PAX9	n/a	No	**	***	(Stockton et al., 2000)
PAX9	n/a	*	**	***	(Frazier-Bowers et al., 2002)
PAX9	n/a	No	*	***	(Nieminen et al., 2001)
PAX9	No	*	**	***	(Jumlongras et al., 2004)
PAX9	n/a	*	**	***	(Goldenberg et al., 2000)
PAX9	n/a	*	**	***	(Arte, 2001)
MSX1	n/a	**	**	*	(van den Boogaard et al., 2000)
MSX1	n/a	**	**	*	(Lidral and Reising, 2002)
MSX1	n/a	No	**	*	(Vastardis et al., 1996)
MSX1	No	No	**	*	(Jumlongras et al., 2001)

Whole blocks of teeth, chiefly incisors and premolars in *AXIN2* cases, premolars and molars in the case of *MSX1*, and molars in the case of *PAX9*, are

found to be missing (oligodontia[2]) in the reported families (Table 5). The clinical picture associated with ectopic canines is much milder and has been termed Incisor-Premolar Hypodontia (IPH).

The genes *MSX1, MSX2, EGF, EGFR, FGF-3* and *FGF-4* have been excluded as candidate genes in humans. (Nieminen et al., 1995, Arte et al., 1996, Mostowska et al., 2003). Major craniofacial and other anomalies are also produced (Jabs et al., 1993, Qiu et al., 1995, Stockton et al., 2000, van den Boogaard et al., 2000). There are no effects reported on the eruption of teeth in affected families. The vast majority of research carried out on tooth agenesis has been on rodents. This causes problems as the dentitions are dissimilar. The incisors, unlike human teeth, are continuously erupting and there is only one dentition in rodents, corresponding to the primary dentition. In humans only the secondary dentition exhibits IPH, the primary being minimally affected. To date, experiments have concentrated on demonstrating the effect on odontogenesis only; no molecular studies have been published showing any association of these genes with tooth eruption. An attempt has been made to establish a link between ectopic canines and the genes *Pax9* and *Msx1* (Peck et al., 2002), however the evidence for this is scanty (Camilleri, 2005). The experimental oligodontia produced in *Msx1* and *Pax9* knockout animals also affects whole blocks of teeth, molars or incisors.

The heterogenic nature of tooth agenesis has made it difficult to identify the culpable genes (Arte, 2001). Much research has been dedicated to identifying the cause of IPH, this being the subject of a COST (European Cooperation in

[2] The term oligodontia is taken to mean agenesis of more than 6 teeth.

Science and Technology) action in the late 1990s. However this has been fruitless to date, so the possibility of using IPH to identify the cause of EC remains remote. Likewise, no mutations have been identified as being responsible for other heritable dental anomalies associated with EC.

1.5 Study Design

1.5.1 Segregation Analysis

Previous studies have indicated that the condition is transmitted in an autosomal dominant fashion (Svinhufvud et al., 1988). The most powerful methods for investigating Mendelian disorders are either exome sequencing or parametric linkage analysis. Problems may have been encountered with exome sequencing as family members share large parts of their genome, resulting in a large number of candidate genes if a single family was used. Also, locus heterogeneity may have confounded the results were unrelated individuals to be used. Several three-generation extended pedigrees had been ascertained. One family had a large number of affected members (8/21 members). This family would have been suitable for either parametric linkage analysis or whole exome sequencing, however parametric linkage analysis was deemed to be preferable as the intial investigation.

1.5.2 Multipoint mapping

Multipoint linkage analysis is more efficient than two-point analysis. The chromosomal order of linked loci can be established and the maximum inheritance information can be deduced from the available marker data. This is

particularly relevant to genotyping with SNPs as these are biallelic and consequently less informative. Multipoint analysis is susceptible to LD and this must be taken into account when using dense SNP arrays. The program MERLIN is an analytical method based on binary gene flow trees (Abecasis et al., 2002). MERLIN can perform both parametric and nonparametric multipoint linkage analysis, using dense marker maps and present this information in the form of haplotypes and whole genome scans. MERLIN can handle more than 100,000 markers on a chromosome when the family is of moderate size. MERLIN can also take LD between markers into account (Abecasis and Wigginton, 2005), though the program encounters problems when analysing small numbers of individuals (Rüschendorf, personal communication). Although the maximum number of members the program can handle is limited (max bits < 17) it can handle larger pedigrees than other linkage analysis programs. A major advantage of MERLIN over other programs is error detection, particularly non-Mendelian inconsistencies. As SNP genotyping errors are often not detected as Mendelian inconsistencies, this is a significant benefit.

1.5.3 Choice of marker

The SNP 6.0 chip has close to a million markers and dedicated copy and a similar number of dedicated Copy Number probes. The number of SNP markers is in excess of requirements, however would be useful were fine mapping to become an option. The present sample size is too small for association analysis to reach significance, however were the sample size to be

increased in the future; the data may well be useful. Copy Number Variation analysis would be carried out, with the advice of Dr. Bert Eussen, Dept. of Clinical Genetics, Erasmus MC, The Netherlands.

Accordingly it was decided to carry out parametric linkage analysis on an informative extended family, using the Affymetrix SNP 6.0 chip for genotyping. The program MERLIN would be used for analysis, with advice from Dr Franz Ruschendorf, Max Delbruck Centre, Berlin.

1.5.4 Whole Exome Sequencing

This would be carried out on two distantly related affected members of the pedigree. Heterozygous, nonsynonymous variants or frameshift indels within suggestive LOD regions would be prioritised as candidate genes. These variants would then be examined by Sanger sequencing in all members of the pedigree to confirm their segregation with the phenotype and also in a number of unrelated individuals.

1.5.5 In situ hybridisation

In situ hybridisation would be carried out on candidate genes in order to determine if they are expressed in developing tooth tissue. The procedure would be carried out in the Department of Craniofacial Development and Orthodontics, King's College London.

Chapter 2 Aims of the Study

2.1 Aims and objectives

The study has two aims. In view of the controversy over the aetiology of the trait, the first aim is to establish that EC does indeed have a genetic aetiology. The objectives here are to determine the familial relative risk, in order to establish that familial clustering is consistent with a genetic component and to determine the most feasible genetic model for the trait.

The second aim is to attempt to identify the gene or genes responsible for EC. The objectives are to carry out a positional cloning experiment in order to identify transmission of a haplotype responsible for the trait. Sequencing of the areas of interest will highlight potential variants and in-situ hybridisation will determine whether the tissue of expression of these variants is appropriate. Unrelated cases will be examined in order to determine whether the same variants are responsible. The genes *MSX1* and *PAX9*, as well as the SNPs identified as being associated with tooth eruption, would be examined for linkage and common variants.

2.2 Null Hypothesis

The null hypothesis is that there is no genetic aetiology for EC.

2.3 Potential Benefits

Identification of a functional mutation may lead to development of a clinical assessment for the condition, whereby affected individuals may be recognised early and the appropriate interceptive measures taken.

The eruption of teeth is dependent on the successful formation of an eruption pathway in the bone, through osteoclastic action (Cahill and Marks, 1980). Identification of the gene product may allow local control of osteoclastic activity, influencing the path of eruption of teeth. This is of particular relevance to the treatment of cleft lip and palate patients, where eruption of teeth into facial bone grafts is a major factor in stabilising the graft, ensuring its long-term success (Bergland et al., 1986).

Control of tooth eruption would reduce or eliminate the need for surgical procedures associated with buried and impacted teeth.

Incisor-premolar hypodontia is closely linked to EC. Identification of the genetic cause of EC may facilitate identification of the genetic cause of hypodontia.

Chapter 3 Materials and Methods

3.1 Segregation analysis

Familial clustering is observed in many EC cases, but this is not in itself evidence of a genetic aetiology, as family members share a similar environment. The weight of evidence is consistent with a genetic aetiology, however a controversy does exist (Becker et al., 1981) and although the mode of transmission is thought to be autosomal dominant (Peck et al., 1994, Svinhufvud et al., 1988) other modes of transmission may exist (Feichtinger et al., 1977). Determination of the genetic component is necessary prior to embarking on a study of this type and this may be carried out by twin studies or by complex segregation analysis. A segregation analysis will provide a genetic model for carrying out linkage or association studies by quantifying the genetic component, determining the mode of transmission, establishing whether the trait is due to a single major gene or the product of several minor genes and determining the allele frequency and penetrance of the gene.

A familial relative risk computation and a segregation analysis would first be carried out in order to determine whether the familial clustering is consistent with a genetic aetiology and also to determine the most feasible genetic model.

Ethical approval was obtained from the University of Malta Medical School Ethics Committee (Ref. 063/2004; copy in Appendix I). Thirty-seven consecutive probands with ectopic maxillary canines were identified during routine clinical examination at a private clinic (SC) and at the School Dental Clinic, Floriana.

The maxillary canines were classified as being ectopic if the tooth was present but eruption had not taken place by the age of 16; if eruption had taken place either palatal to the line of the arch or high in the buccal sulcus, above keratinsed gingiva or was transposed or pseudotransposed; and crowding should not be the major factor in the position of eruption of the tooth. In the case of a congenitally missing lateral incisor, one or both of the canines had erupted mesial to its normal position, in place of the missing lateral.

The inclusion criteria were Maltese ethnicity, taken as both sides of the family having been resident for a minimum of three generations and a positive history of ectopic (buccally or palatally displaced) canines or failure of eruption of the canine tooth by the age of 16. Individuals affected by an orofacial cleft or a genetic syndrome likely to influence eruption of teeth were excluded. A further set of twenty-six consecutive families was selected on the strength of two probands, in order to identify families with a stronger genetic predisposition. Families with an affected twin were investigated even if the other twin was unaffected. Ectopic or missing teeth were identified by clinical examination, from existing records or from the family history. Radiographic investigation was undertaken only where clinically necessary. The pedigrees may be found in Appendix 2.

3.1.1 Data analysis

The proportion of affected relatives in different groups was compared using a chi square test or Fisher’s exact test, where appropriate. The pedigree data was

analysed to assess the familial risk of EC and other related phenotypes, using the familial relative risk computed.

The segregation analysis program Pedigree Analysis Program PAP, (Hasstedt, 2005) was used to model the inheritance of EC throughout the 63 pedigrees. Pedigree members were defined as affected or unaffected with EC (with no further clinical information used). Genetic models assumed a single locus model with penetrances and mutation frequency maximized for each model. Nested models were compared using the likelihood ratio test, where the difference in -2log (likelihood) has a chi-square distribution, with the degrees of freedom equal to the difference in the number of additional parameters fitted in the most general model. Sporadic (no genetic effect), recessive and dominant models were compared to a codominant model. A polygenic model was also tested and compared to the dominant model. An ascertainment correction for the probands in each family was applied (Khoury et al., 1993).

Five sets of monozygotic and six sets of dizygotic twins were included in the sample. A further two sets of monozygotic twins and one set of triplets were subsequently referred by a colleague aware of the research focus in twins. These were not included in the pedigree or segregation analyses however were used in assessing the pairwise concordance.

3.2 Linkage analysis

There have been few linkage studies on tooth eruption disorders. Cherubism is a familial autosomal dominant disorder that may be related to tooth development and eruption. A study by Tiziani established a linkage to

Chromosome 4p.16 with a LOD score of 4.21(Tiziani et al., 1999). Mangion performed a genome wide linkage search and localized the gene to chromosome 4p16.3, with a LOD score of 5.64. (Mangion et al., 1999). Using linkage and haplotype analysis, Ueki localised the area to a 1.5Mb region and detected point mutations in the *SH3BP2* gene (Ueki et al., 2001). Parametric linkage analysis carried out by Decker on a 3-generation family found splice site mutations in the *PTHR1* gene in affected individuals. These mutations are thought to cause premature degradation of the precursor or nullify the function of the receptor, causing haploinsufficiency (Decker et al., 2008). Other studies have since confirmed *PTHR1* as the causal gene for Primary Failure of Eruption (Yamaguchi et al., 2011). In genome wide association studies, Pillas et al found five loci with suggestive association to the timing of primary tooth eruption (Pillas et al., 2010). Geller et al found four loci associated at genome wide significance with the timing of permanent tooth eruption in subjects between the ages of 6 and 14. One SNP was also associated with the timing of deciduous tooth eruption (Geller et al., 2011).

An informative, three-generation, 21 member family with eight members exhibiting EC had been identified (Figure 4). The phenotypes were detailed (Figure 5).

The phenotype of individual 41 was different from the rest of the affected kindred in that his canine had erupted high and buccal, though crowding was not evident. In addition, he did not exhibit hypodontia and his lateral incisors were of normal size. Provision was made for his being a phenocopy.

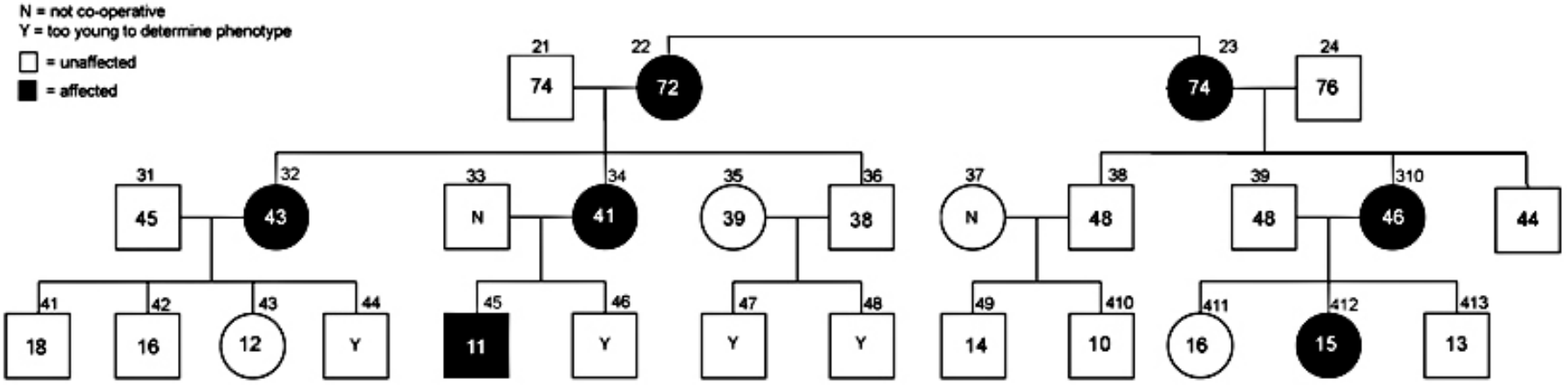


Figure 4. The family tree used in the study. The pedigree numbers are outside the symbol, the age of the subject is inside. Males are square symbols, females round. Individual #41 was categorised as a phenocopy and so is not marked as affected.

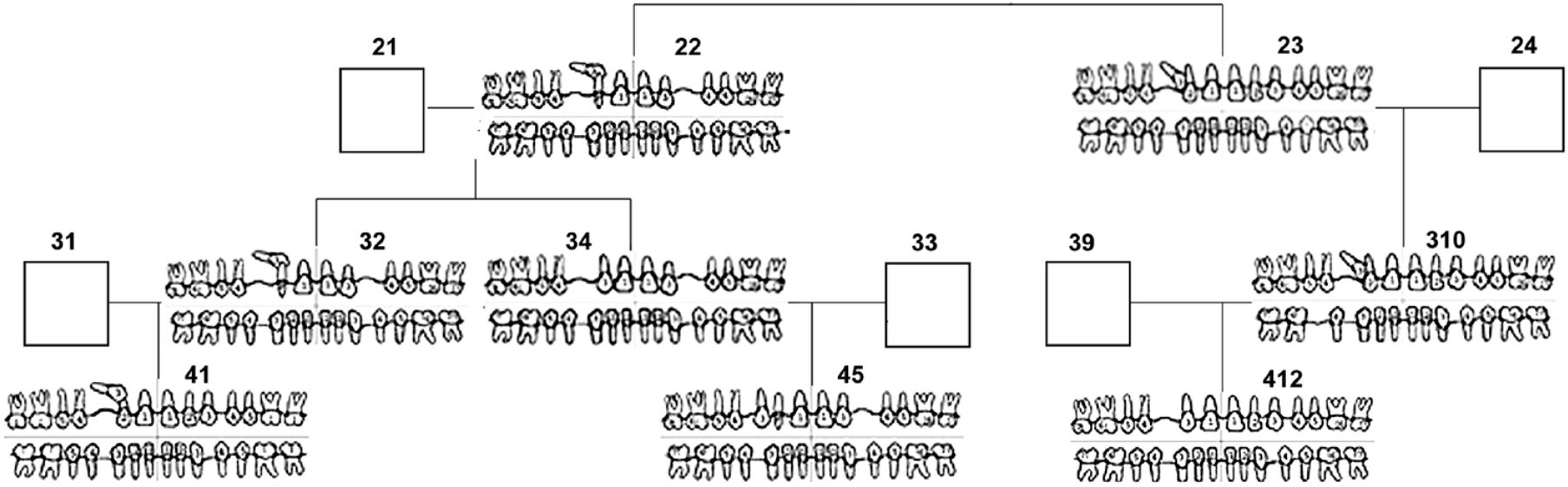


Figure 5. The phenotypes of the affected members of the family. EC was present in all members. IPH was present in all affected members except #23 and #41. The canine in #41 was buccal and this individual was categorised as a phenocopy.

Genomic DNA was obtained from saliva using Oragene kits (DNA Genotek, CA) and extracted according to the recommended protocol.

Equipment: Water bath set at 50°C

Reagents: Oragene·DNA Purifier (OG-L2P, Ethanol (70% and 90%) at room temperature, TE buffer (10 mMTris-HCI, 1mM EDTA, pH 8.0) microcentrifuge capable of running at 13,000 rpm (15,000 x *g)*

The entire Oragene·DNA/saliva sample was mixed by inversion and gentle shaking for a few seconds. It was incubated in the water bath at 50°C for over 1 hour. 500µL of the sample was transferred to a 1.5 mL microcentrifuge tube and 20 µL ($1/25^{th}$volume) of Oragene·DNA Purifier OG-L2P added to the microcentrifuge tube and mixed by vortexing for a few seconds. It was then incubated on ice for 10 minutes and centrifuged at room temperature for 5 minutes at 13,000 rpm (15,000 x g). The clear supernatant was transferred with a pipette tip into a fresh microcentrifuge tube and the pellet containing impurities discarded. To 500 µL of supernatant, 500 µL(i.e., an equal volume) of room- temperature, 90% ethanol was added and mixed gently by inversion 10 times. The tube was left to stand for 10 minutes at room temperature then placed the tube in the microcentrifuge in a known orientation and centrifuged at room temperature for 2 minutes at 13,000 rpm (15,000 X g). The supernatant was carefully removed and 250 µL of 70% ethanol was added and left to stand at room temperature for 1 minute. The ethanol was then carefully removed without disturbing the pellet. The 70% ethanol wash was repeated, then all ethanol was removed and the pellet dried in air overnight. The pellet was then

dissolved in 100 µL TE buffer by incubating at room temperature for at least 8 hours, aliquoted and stored at -20°C for a maximum of three months.

The DNA was further purified using the phenol-chloroform method.

Saturation of phenol with water

Reagents: Phenol ,Water Bath, 8-hydroxyquinoline, Chlorofom, Isoamyl alcohol.

Equipment: Large separating funnel, measuring flasks, Pasteur pipette, pH meter (Metroohm654 pH meter, Switzerland), Fume cupboard,

Method: The crystalline phenol was placed in a water bath at 65°C until melted and twice the volume of distilled water was added to the phenol in a large separating funnel, in the fume cupboard. The phenol and water were mixed to a fine emulsion and then allowed to separate and the lower organic layer was collected. This step was repeated until the pH of the upper aqueous phase was found to be between pH3.0 and 4.5. 8-hydroxyquinoline was added as an antioxidant.

An equal volume of 200mM Tris-HCl was added to water-saturated phenol in the separating funnel in the fume cupboard. The two phases were mixed well, allowed to separate at room temperature and the lower organic phase was removed. This step was repeated another two times. The equilibrated phenol was aliquoted and stored in light-tight containers under a layer of 100mM Tris-HCl (pH 8.0) at -20°C indefinitely.

Chloroform was added to isoamyl alcohol in the proportion of 24:1, then equal parts of phenol and chloroform:isoamyl alcohol were mixed. The mixture was

stored in light-tight containers under a layer of 100mM Tris-HCl (pH 8.0) at 4°C for up to one month.

The molecular weight of the DNA was assessed by visualisation on 0..5% agarose gel.

Preparation of 50X TAE running Buffer

Equipment: 1L Beaker, magnetic stirrer.

Reagents: Tris-base: 242 g, Acetate (100% acetic acid): 57.1 ml, EDTA: 100 ml 0.5M sodium EDTA.

Method: The ingredients were mixed and distilled water added up to one litre. The stock mixture was stored at 4C. It was diluted for use by addition of 20 mL solution to 980mL distilled water.

Preparation of gel

Reagents: Agarose Gel, TAE running buffer, distilled water, Ethidium bromide, sealing tape.

Equipment: Pyrex beaker, gel box and comb, microwave oven.

Method:The TAE buffer was diluted to 1X and the appropriate mass of agarose measured out into a beaker with the appropriate volume of buffer. In this case, 100mL was sufficient for most purposes. The beaker was microwaved until the agrose was fully melted. While the agarose was cooling, the open edges of the gel box were sealed with tape. The agarose solution was poured into the taped gelbox. and left to cool for about 30 minutes, until the gel was solid.

The molecular weight of the samples was determined by electrophoresis.

Materials: Mixing dish, electrophoresis bath, UV hood connected to scanner.

Reagents: Loading buffer, 100 base-pair ladder.

Method: The comb was removed from the gel and placed in the electrophoresis bath. 4µL template DNA/PCR product was placed in the wells of the agarose gel. One well was reserved for the 100 base-pair ladder. The voltage was set to 120v and left for 15 minutes. The gel was removed from the bath, placed under the UV hood and scanned. The images were saved to disk or discarded as appropriate.

Purity was determined by spectrophotometry (Biophotometer, Eppendorf AG, Germany).

Equipment: Tubes with a minimum capacity of 50µL, Spectrophotometer (Biophotometer, Eppendorf AG).

Reagents: Distilled water or TE buffer

Method

The spectrograph was set to read for double stranded DNA. 5µL of DNA solution was diluted in 45µL (10x dilution) TE buffer. The spectrophotometer was calibrated to zero using TE buffer, noting the orientation of the cuvette. The cuvette was rinsed thoroughly using distilled water and dried. 50µL of the diluted DNA was placed in the cuvette, with the same orientation as the blank and the concentration and the 260/280 reading noted. Recalibration was carried out every 10 readings.

3.2.1 Genotyping

This was carried out at the Erasmus-MC laboratories in the University of Rotterdam (Netherlands Consortium for Systems Biology, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands) using Affymetrix Genome-Wide

Human SNP Array 6.0 chips. These chips contain 1.8 million genetic markers, including over 906,600 SNPs spaced at 1-5 kilobase (Kb) intervals throughout the genome and more than 946,000 probes for the detection of CNVs (Table 3).

3.2.2 Quality Control

Genotypes were called using the Birdseed Algorithm for Genome-Wide Human SNP Array 6.0. The quality control criteria for the genotyping data were set at a sample call rate of over 90%. Quality control call rate ranged between 90 and 98% and MAPD values were within bounds.

3.2.3 Compilation of pedigrees

The full pedigree, named V1, was used for quality control. The LOD score of the pedigree was estimated by assigning a score of 0.3 for each informative meiosis and subtracting 0.6 where the phase of the parents was unknown. Using this method, the maximum LOD score extractable from V1 was calculated to be 3.9. Assuming autosomal dominant inheritance and full penetrance, individuals #37, #49 and #410 would not contribute towards the final LOD score and would add to the computational burden.V1 was modified by removal of these individuals and renamed V2. The large family was split into two in version V3 in order to include all individuals. Individuals #37, #410, #49, #43 and #413 were removed from V1 in order to increase the speed of the run and used for quality control (V4).

Although the criteria for ectopic eruption were satisfied, the phenotype of individual 41 differed from the rest of the affecteds in the family. In order to

avoid the possibility of phenocopies, the dissimilarity warranted his exclusion. #41 was coded unknown (V6) and unaffected (V7). The maximum LOD score from pedigree V7 was 3.9.

As the segregation analysis and quality control checks indicated reduced penetrance, a parametric 'affected only' analysis was also carried out. All unaffected members were coded as unknown, this pedigree being named V8. The maximum LOD score for V8 is 1.5. These pedigrees may be found in Appendix 2. Pedigrees V8 and V7 were used in the final analysis.

3.2.4 Data analysis

The program Alohomora (Ruschendorf and Nurnberg, 2005) was used to convert Affymetrix genotype data into linkage and haplotype information, create Merlin input files and to carry out quality control. The Caucasian allele frequency file "GW6_freqA_Caucasian_na23" and Genetic map "deCodeGW6_map_decode_na30_2009-11-13" were used (Affymetrix, USA). The linkage analysis program MERLIN (Abecasis et al., 2002), was used to carry out parametric multipoint linkage analysis on the pedigrees. The program was run at the Max Delbruck Centre, Berlin.

MERLIN gives the option to use a sliding window to analyse the data, reducing computational requirements, however as this may result in false positive peaks at the peripheries of the window (Rüschendorf, personal communication), this option was not used in the pedigree analyses and each chromosome was calculated as a whole.

As the degree of penetrance was unknown, MERLIN was used to run a conservative, parametric 'affected only' analysis with the pruned marker set on pedigree V8 using a dominant model and a minor allele frequency (MAF) of 0.001. Further analyses were also run on pedigree V7 with the genetic model: dominant with reduced penetrance, MAF 0.001, penetrances AA=0.00, AB= 0.90, BB= 0.90, where A is the unaffected allele and B is the minor, disease allele.

3.2.5 Genotyping Quality Control

In order to verify the relationship between individuals, the data were subjected to standard quality assurance checks. Gender checking and relationship checking was carried out using the Graphical Representation of Relationship (GRR) program (Abecasis et al., 2001) and checking for Mendelian errors (PedCheck) (O'Connell and Weeks, 1998) were performed through Alohomora on pedigree V1. "Unlikely genotypes" are equivalent to double recombinations in a short chromosomal segment and may be due to genotyping errors or a wrong SNP position in the genetic map. These were deleted using Alohomora.

3.2.6 Linkage Disequilibrium

In order to exclude the effects of LD, two calculations were run using pedigree V4. An autosomal dominant model with full penetrance was assumed, with inter-marker distances of 10 Kb and of 50 Kb. A discrepancy between the results raised the suspicion that LD may be in operation. Merlin does take LD between markers into account, however large numbers of subjects are

necessary, program operation being problematical with one single family (Rüschendorf, personal communication). Further testing involved the use of a control set of 1460 samples, also genotyped with the Affymetrix SNP 6.0 chip. Here, markers were tested for Linkage Disequilibrium using the program PLINK (Purcell et al., 2007) and the option " --indep-pairwise 1500 150 0.2", where 1500 is the size of the moving window, 150 is the size of the window shift and 0.2 is the maximum value of r^2 allowed. A pruned set of 86542 SNPs genome-wide was created, where the value of r^2 did not exceed 0.2.

3.2.7 Haplotyping

This was carried out on the areas of interest using Haplopainter (Thiele and Nurnberg, 2005).

3.2.8 Exclusion of linkage

Using data derived from Pedigree V8, the LOD scores of markers spanning the genes, *SH3BP2, PTHR1, MSX1* and *PAX9* were examined in order to exclude the possibility of linkage. SNPs flanking those identified by Geller et. al. as being associated with the timing of permanent tooth eruption were similarly treated.

3.2.9 Copy Number Variations

The affected members of pedigree V8 were selected and allocated to two groups according to family branch. The unaffected members formed the control group. The data were then analysed using Nexus Copy Number

(Biodiscovery, USA) by comparing the average numbers of events in each affected group first separately and then in combination, against the control group.

3.3 Whole Exome Sequencing

3.3.1 Exome capture and library preparation

This was carried out at the Department of Genetics, King's College London. DNA from individuals 32 and 410 used as these samples were readily available. Sample preparation for exome sequencing was undertaken using the Agilent SureSelect All Exon Target Enrichment System for Illumina Paired-End Sequencing. Briefly, 3µg of genomic DNA were diluted in 120µl TE and sheared into 150-200 bp DNA fragments, using a Covaris S220 ultra-sonicator (Covaris, USA) according to the manufacturer's instructions. DNA fragments were purified using the Agencourt AMPure XP beads **(**Beckman Coulter International SA**)** on a Dynal magnetic separator (Invitrogen Corporation) and their size was assessed on an Agilent 2100 Bioanalyzer (Agilent Technologies UK), using DNA 1000 chips. Next, DNA end-repair, 3'-dA overhang and adapter ligation were carried out using the NEBNext® DNA Sample Prep Reagent Set 1 (New England Biolabs, USA), according to the manufacturer's instructions. The necessary purification steps were carried out using Agencourt AMPure XP beads and the size of the DNA fragments was again assessed on an Agilent 2100 Bioanalyzer.

The adapter-ligated libraries thus generated were amplified by PCR using the Herculase II Fusion DNA Polymerase Kit (Agilent Technologies UK) and

purified using the Agencourt AMPure XP beads. The libraries were quantified on a Qubit Fluorometer (Invitrogen Corporation, USA) and analysed with the Agilent 2100 Bioanalyzer. Five-hundred and fifty nanograms of each library was concentrated in a 3.4µl final volume and after 24h of hybridization with the target RNA baits, libraries were captured with Dynal MyOne Streptavidin T1 magnetic beads. The hybridised libraries were purified on a Dynal magnetic separator and amplified using the Herculase II Fusion DNA Polymerase Kit. The amplified and captured libraries were purified with Agencourt AMPure XP beads, assessed with an Agilent 2100 Bioanalyzer, using the 2100 High Sensitivity chip and quantified with a Qubit Fluorometer using the HS dsDNA assay kit. Based on the average length of DNA fragments, libraries were diluted to a 1nM final concentration, in a volume of 32µl, according to the following formula:

$$\frac{\mu g\ DNA}{660pg} \times \frac{10^6 pg}{1\mu g} \times \frac{1}{N} = pmol\ DNA$$

where N is the length of the fragment and 660pg/pmol is the average molecular weight of a nucleotide pair.

3.3.2 Massive parallel sequencing

Clonal cluster generation was performed in a cBot Cluster Generation System (Illumina Inc, USA), using the Standard Cluster Generation Kit v4 according to the manufacturer's guidelines. Initially, single–molecule DNA templates were immobilised in the flow cell and copied by 3' extension. After removal of the original DNA templates, the immobilised copies were subjected to isothermal bridge amplification. Following linearisation, blocking and primer

hybridisation of the DNA clusters, 76bp paired end reads were generated in a Genome Analyser IIx (IlluminaInc, USA), using the TruSeq SBS Kit v5–GA Kit.

3.3.3 Stepwise filtering approach for exome sequencing

Exome sequencing was carried out using the SureSelect All Exon Target Enrichment System (Agilent Technologies UK) and sequencing on a Genome Analyser IIx (IlluminaInc USA) with 76bp paired end reads. Sequence reads were aligned to the human reference sequence hg18 using the Novalignsoftware (Novocraft Technologies SdnBhD). PCR and optical duplicates were removed with MarkDuplicates (http://picard.sourceforge.net/index.shtml) and poor quality alignments were removed using the SAMtools software (Li et al., 2009).Capture efficiency and completeness were assessed with the BEDtools package (Quinlan and Hall, 2010).Variants with a minimum 4X coverage were called using SAMtools and each call set was evaluated based on the number of variant sites and the transition to transversion mutation ratio. Variant annotation was carried out using the Annovar software (Wang et al., 2010). Variants that were not present in either dbSNP (build 129 and 132) or the 1000 genomes dataset (December 2010 data release) or in 250 additional control exomes available in-house were considered as novel. The data was analysed using Microsoft Excel, filtering for heterozygous, nonsynonymous variants or frameshift indels with a frequency of < 0.001 in the 1000 Genomes or dbSNP databases. Variants were also checked against the pooled allele frequencies of 50 Maltese subjects. Variants

present at a frequency < 0.1were also included in subsequent analysis. Variants not present in both sets of exomes were then discarded. Variants that were not consistent with the inheritance pattern of each trait and with prior linkage data were filtered out. As the disease has been predicted to be dominant, homozygous variants were discarded. Nonsynonymous SNVs or frameshift insertions/deletions (indels) in exons or splice sites were prioritised. Pathogenicity of putative disease causing variants was assessed using the SIFT algorithm (Kumar et al., 2009) and PolyPhen2 (Adzhubei et al., 2010). Gene expression pattern and function, where known, were also taken into account in selecting candidates for further investigation.

DNA was also collected from 18 other unrelated EC cases recruited consecutively, for validation of the results.

3.3.4 Sanger Sequencing

The four variants were sequenced in the whole family, Sequencing of exon 1 and intron-exon boundaries of *PPP1R14C* as well as exon 17 of *ANO5* was carried out on 18 other unrelated sporadic cases. Details of the primers and thermocycler programs used for PCR are given in Tables 6 and 7. Reddymix 1x mastermix (Thermo Scientific, USA), was used following the manufacturer's instructions. PCR products were visualised on 1.2% agarose gel.

Table 6. Primers for Investigation of Candidate Genes.

	Primer Name	Primer Sequence	Annealing Tempera-ture	Product Size (nt)
PPP1R14C	PPP1F	GTCCTCGGCGGCTTCTTT	57°C	579
	PPP1R	GGGCACAGAGGAGCTAGAGA		
ANO5	ANO5F	AACCCTTCCAACCAAAACCT	52°C	190
	ANO5R	AAAAATCTGTTTCCCGGTCA		
EPM2A	EPM2AF	TCCTACCGTTCTGTTTTTCACTC	55°C	434
	EPM2AR	CTGCAGTTTCGAACACAGTAAA		
VSTM2L	VSTM2LF	CAGCTGAAGCCAGACCACTC	56°C	482
	VSTM2LR	AGGAAGAAGGGAGGGAGACA		

Table 7. Conditions for PCR. The same program was used for all four primer pairs.

Cycle		
95°C	5 min	
95°C	20 seconds	30 cycles
**°C	35 seconds	
72°C	50 seconds	
72°C	7 minutes	
5°C	hold	

3.4 In situ hybridisation

In situ hybridisation was carried out in the Department of Craniofacial Development and Orthodontics, Dental Institute, King's College on human embryos. In situ hybridisation (ISH) was carried out at the epithelial thickening

stage and the bud, cap and bell stage for both PPP1R14C and ANO5.

The probe, chamber solution and hybridisation buffer were prepared.

Synthesis of 32S Probe

Mix:

5X reaction buffer 5.0µl:

1M DTT 0.5µl

10mM GTP 1.2µl

10mM ATP 1.2µl

10mM CTP 1.2µl

50µM UTP 1.0µl

DNA 6.4µl

RNasein (40U/µl) 0.5µl

35S UTP 7.0µl

Polymerase enzyme 1.0µl

and incubate 40 mins at 37°C. Add 1µl Polymerase enzyme more and incubate 60 mins at 37°C. Add:

DNase 2.5µl /tube

RNase Inhibitor 0.5µl

tRNA (10mg/ml) 1.0µl

1M DTT 0.5µl

RNase-Free DNase 0.5µl

Incubate 10 min at 37°C . Then add

DEPC H2O 160µl

1M DTT 4.0µl

5M NaCl 4.0µl

3M NaOAc 20µl (Sodium Acetate) pH 5.2

EtOH 400µl

Precipitate O/N at -20C, 588µl/tube, or 1h at -80°C, centrifuge at 13,000rpm for 10mins and add 10mM DTT/EtOH Wash 500µl/tube, centrifuge at 13,000rpm for 10mins and add 50mM DTT/H2O 50µl/tube. Add Hydrolysis Solution 50µl/tube at 60℃. Add 10mM DTT/EtOH Wash 500µl/tube, centrifuge at 13,000rpm for 10mins and add 50mM DTT/H2O 50µl/tube. Add Hydrolysis Solution 50µl/tube at 60℃

Add Neutralising Buffer 50µl/ tube (on ice), add PPE 504µl/tube. Place in -80°C freezer for 1 hour, centrifuge at 13,000rpm for10mins and add 10mM DTT/ EtOH wash 500µl/tube. Centrifuge at 13,000rpm 5mins and dry for 5 mins. Add 50mM DTT/ H2O 50µl / tube and store in -80°C freezer.

Solution for Chamber 100 ml/1 Chamber

10 mls 20X SSC

50 mls Formamide

40 mls H2O

Hybridisation Buffer

885µl = Buffer for 1ml (final probe)

Formamide 500µl

50% Dextran Sulfate 200µl

5M NaCl 60µl

20mg/ml tRNA 25µl

1M DTT 50µl

1M Tri-Hcl pH8.0 20µl

0.5M EDTA 10µl

50X Denhalt's 20µl

(2 X 10^4/cpm) X number of slides X 100µl = probe for Hybridisation buffer

S35 probe + Hybridisation Buffer + DEPC H2O = 100µl for 1 slide

***E.g**. if cpm=800,000 and 4 slides for in situ, (2 X 10^4/800000) X 400 = 10*

S35 probe 10µl + Hybridisation Buffer 354µl + DEPC H2O 36µl = total 400µl

The tissue was frozen, sectioned, dehydrated and fixed with paraformaldehyde. It was then permeabilised with Proteinase K, to allow penetration of the probe in order that it may hybridise with the target. The ^{32}S probe was mixed with hybridisation buffer. The hybridisation solution was then applied and heated to denature the probe. Excess probe was digested with RNase and washed off. Dithiothreitol and Ammonium Acetate were added. The hybridisation was visualised by dipping in radiosensitive emulsion and prolonged darkroom exposure (2 weeks). The slides were then developed and fixed.

Day1

Place slides in 15 mins Histoclear X2, then place in 2 mins 100% EtOH. Rehydrate by placing in

2 mins 95% EtOH (17.5 mls H2O + 332.5 mls EtOH)

2 mins 90% EtOH (35 mls H2O + 315 mls EtOH)

2 mins 80% EtOH (70 mls H2O + 280 mls EtOH)

2 mins 60% EtOH (140 mls H2O + 210 mls EtOH)

2 mins 30% EtOH (280 mls H2O + 70 mls EtOH)

2 mins DEPC H2O (repeat)

20 mins HCL in DEPC H2O

Permeabilise by digestion with Proteinase K by placing for

5 mins 2 X SSC (35 mls 20X SSC + 315 mls H2O)

10 mins ProK (35 mls 1M Tris pH8)

2 mins 2mg/ml Glycine in DEPC PBS (700µgs Glycine+350 mls PBS)

1 min PBS (Repeat step)

Fix for 20 mins in 4%PFA in PBS (14 g PFA + 200 mls PBS)

Rinse for 2 mins PBS

Acetylate by placing for 10 mins 0.1M Triethanolamine (TEA) in H2O (4.65mls TEA +345.35 mls H2O+ HCL 800µl +875µl acetic anhydride (AA)) Stir thoroughly

Wash for 5 mins with PBS followed by 2 mins in DEPC H2O

Place for 2 mins 30% EtOH

2 mins 60% EtOH

2 mins 80% EtOH

2 mins 90% EtOH

2 mins 95% EtOH

2 mins 100% EtOH

Day2

Prepare 2L Wash solution mix 200 ml 20X SSC with 1L Formamide, 800 ml H2O and g DTT

Rinse slides for 15 mins in Wash Solution at 55 °C, followed by a further 20 mins rinse at 55 °C and a final rinse at 65 °C. Then rinse twice for 15 mins rinse in RNase Buffer at 37 °C

Rinse for 30 mins in RNase A (40µg/ml = 10 mg/250 ml) at 37 °C followed by a further 15 mins rinse in RNase Buffer 37 °C and two rinses of 20 mins rinse in Wash Solution at65 °C. Rinse for 20 mins in 0.1 X SSC/DTT at 65 °C and a further 5 mins rinse in 0.1 X SSC/DTT RT.

Rinse for 2 mins in Ammonium Acetate/ 70% EtOH, 2mins in 95% EtOH, 2mins in 100% EtOH and leave for 1 hour to dry

Day3 Dipping

Waterbath 45 °C, 2% Glycerol, 1:1 emulsion

After dipping, leave slides 2 hours in dark box, then slides should be left in 4°C.

2 weeks later. Developing

Developer:

Developer: 31.6 g/200 ml dH2O. Warm water up to 40°C. After dissolving powder, cool to 17°C.

Fix: Fix 50ml, dH2O 150 ml

Place slides in developer 5mins and wash with tap water. Fix for 5mins and wash with running tap water for 1hour. Dip slides in distilled water and place in Hematoxylin (1/4) for 2-4 mins. Wash with tap water for 10mins

Control slides detect false positives resulting from background, either by targeting another area of the transcript or by using unlabeled or antisense probes. In this situation, due to the scarcity and expense of the tissue sections, control slides were not specifically used, however both ISH procedures were carried out on adjacent slides, with the different slides acting as each other's control.

Chapter 4 Results

4.1 Segregation analysis

Sixty-three families were identified in total. The number of individuals of known dental status was 524. One hundred and thirty-nine individuals had EC. The percentage of dental anomalies in the sample for first degree relatives was noted and differences against published population prevalences investigated.

In the first degree relatives of EC probands there was a significantly higher prevalence of EC (15%, $p < 0.001$) and lateral incisor agenesis (7.88%, $p = 0.01$) as compared to figures for the Maltese population.

There were eight cases of maxillary canine transposition in the whole sample. Three were probands, five were first degree relatives; one was a second degree relative. Seven were in the upper jaw, giving a prevalence of 1.4%. This is significantly higher ($p<0.001$) than the prevalence estimated in a Caucasian population. Two cases of mandibular canine-lateral incisor transposition were also recorded.

The percentage of first, second and third degree relatives with ectopic maxillary canines was used to calculate the relative risk ($\lambda_{R)}$ to the first, second and third degree relatives. This is calculated as ($\lambda_R = \kappa_R/\kappa$) where κ is the population prevalence and κ_R is the percentage of relatives affected according to the type of relative i.e. of the first, second or third degree. Ascertainment bias was corrected for by exclusion of the probands from the calculation. The result is shown in Table 8. The relative risk was then used to plot the dropoff for each decreasing degree of unilineal relationship (Figure 6). There was no

difference in EC risk between families ascertained from one or two probands for first, second or third degree relatives ($p = 0.49$, $p = 0.52$, $p = 0.65$). Nor was there a difference in the numbers of sib-sib and parent-offspring affected pairs for each type of family ascertainment ($p = 0.89$)

Table 8. The proportions of ectopic canine-affected subjects used in calculating the relative risks.

	Relatives affected with ectopic canines		**Total number of individuals**	**Percentage of relatives affected**	**Relative Risk**
	2 probands	1 proband			
1st Degree Relatives	16	15	203	15.27	2.78
2nd Degree Relatives	6	8	117	11.97	2.18
3rd Degree Relatives	5	3	85	9.41	1.71
Average Sampled Family size	10	7			

There was an appreciable gender bias, with the affected Female to Male Ratio for 1^{st}, 2^{nd} and 3^{rd} degree relatives being 1.78. The gender ratio for the probands was 1.64 and elimination of the probands from the calculation gave a ratio of 1.89. However there was no appreciable risk difference between relatives of male or female probands, $p = 0.77$

The proportion of male first degree relatives affected vs. proportion of female first degree relatives was not significant, ($p= 0.12$)

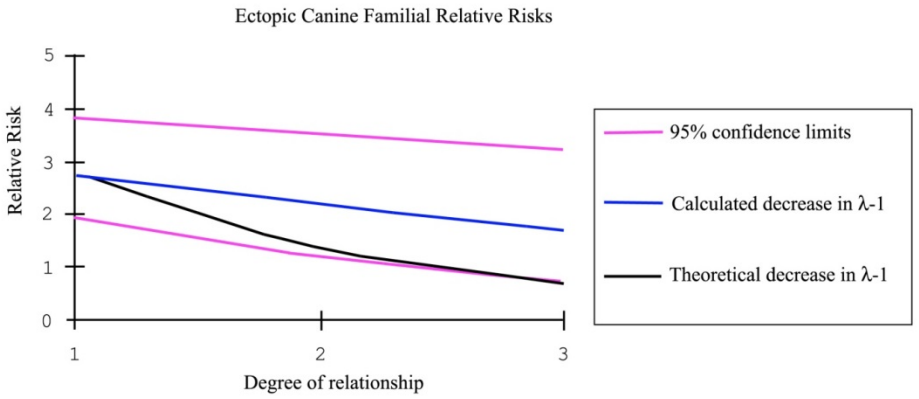


Figure 6. Relative risk dropoff as a function of degree of relation. In diseases where the parameter λR -1 decreases by a factor of 2 for each decreasing degree of unilineal relationship, this is suggestive of a single locus. (λR = Relative Risk).

Although there was no evidence of sex-linked transmission, 85% of the three-generation families show instances where an obligate carrier showed a normal phenotype, although the condition was transmitted to their children. Repeat-sequence diseases exhibit anticipation, i.e. increasing severity or earlier incidence of the disease through generations. There was no apparent pattern of anticipation no augmentation or attenuation of phenotype through the generations was noted.

The segregation analysis program PAP was used to fit different genetic models to the pedigrees. A likelihood ratio test was applied to determine any significant difference in the fit of the various models (Table 9). All the genetic models fitted provided a significantly better fit than the sporadic model, confirming that EC has a genetic basis. The parameter estimates from the co-dominant model were very similar to the dominant model. The dominant model

had the least parameters and therefore provided the more parsimonious model, with the -2log (likelihood) for the recessive model being substantially higher. A mixed dominant model did not provide an improved fit, implying that there is no evidence for a polygenic component in addition to the Mendelian locus. These results showed that a single dominant gene best accounts for the inheritance of EC in these pedigrees, with a mutation frequency of 11%, and penetrances of 36% in mutation carriers. From the dominant model fitted in the segregation analysis, the risk to the sibling of an EC case was approximately 0.18. This compared well with the average estimate of 0.15 obtained from the pedigree analysis. The population prevalence predicted by the dominant model was determined using the formula: $q^2*p1 + 2q(1-q)*p2 + (1-q)^2*p3$, where q is the allele frequency, and p1, p2, p3 the penetrances for the three genotypes. The calculated prevalence was 7%, which also compared well with the estimated population prevalence of 5.5%. This supports the dominant model as being the most likely.

Seven pairs of monozygotic twins formed part of the sample (Table 10). Of these, five were completely discordant, in that one twin was unaffected. Of the other two pairs, one pair was a mirror-image and the other set had one twin affected unilaterally, the other bilaterally. The pairwise concordance in this sample was 28.6%. Curiously, one set of twins exhibited both ectopic and supernumerary teeth, being discordant for the numbers of both ectopic and supernumerary teeth. Seven pairs of dizygotic twins are also on record, with two showing concordance for EC, a pairwise concordance rate of 28.6%.

Table 9. The -2 log likelihoods of the various models tested. The wild-type allele frequency and the penetrance of the genotypes AA, AB and BB were computed for the various models. Chi square tests were used to compare models with a difference of no more than one degree of freedom.

Model	**Allele freq**	**Penetrance**			**Herita-bility**	**-2 log likely-hood**	**No.of Param-eters**	**p-value**	
		AA	AB	BB	h2				
Mixed Dominant	0.89	0	0.32		0.47	379.76	4	-	0.09
Polygenic	-	0.12			0.4	384.49	2	0.07	
Sporadic	-	0.18			-	387.86	1		0.01
Dominant	0.89	0	0.36		-	380.34	3	1.00	
Co-dominant	0.89	0	0.36	0.36	-	380.34	4		0.09
Recessive	0.38	0		0.32	-	383.22	3	-	

Table 10. The pairwise concordance rates of the monozygotic and dizygotic twins/triplets in the sample.

Monozygotic Twins				Multiple births			
	Sex	Concordance (No. of impacted canines)		Sex	Concordance (No. of impacted canines)		
1	F	1	1	F/F	1	2	
2	F	1	0	F/F	1	1	
3	M	1	0	F/F	0	1	
4	M	1	0	F/F	3	0	
5	M	2	0	F/F	2	0	
6	F	1	0	M/M/M	0	1	0
7	F	2	1	F/F/F	0	0	1
	Concordance = 2/7; 28.6%			Concordance = 2/7; 28.6%			

4.1.1 Discussion

4.1.1.1 Dental Anomalies

There was a significantly higher prevalence of EC in first degree relatives of EC probands (15%, $p < 0.001$) as compared to figures for the Maltese population (Camilleri, 1995)(Camilleri et al., 1995 unpublished observations). Familial clustering is a common feature of the trait; however this could be due to environmental as well as genetic factors.

The prevalence of canine transposition in the sample was far higher than the 0.27% prevalence estimated in a Caucasian population (Yilmaz et al., 2005). The highly significant difference is indicative of transposition being an extreme variant of EC and is consistent with the highly variable expression exhibited.

The prevalence of lateral incisor agenesis in the Maltese population is 3% (Camilleri and Mulligan, 2007). The high figure found in association with EC suggests that both traits are related, in keeping with previous evidence (Becker et al., 1981, Svinhufvud et al., 1988, Peck et al., 1994, Pirinen et al., 1996, Baccetti, 1998). A limitation is that the prevalence of premolar agenesis was not calculated as this is difficult to estimate in family members unaffected by EC without undertaking radiographic investigation. As the manifestations of hypodontia are extremely variable, the prevalence may have been underestimated here.

4.1.1.2 Pedigree Analysis

The pattern of transmission was that of an autosomal dominant trait. Expression was highly variable both within and between families. Incomplete

penetrance was evident with obligate carriers unaffected by EC. This may be explained by the highly variable expressivity seen, with failure of penetrance in some cases.

Considering the high population prevalence, the figure for λ_R suggests a major genetic component. When the dropoff in λ_{R-1} (decreasing degrees of relation) was plotted as a function of the relative risk, the gradual decrease for every decreasing degree of relationship was strongly suggestive of a single locus, though it did not rule out genetic heterogeneity (Risch, 1990) (Figure 6). However, there was no difference in prevalence between families ascertained from one or two probands for first, second or third degree relatives. Nor was there a difference in the sib-sib/parent-offspring relationships in the two types of family. This is indicative of a single gene as multiple genes would be expected to give a higher prevalence in families ascertained from two probands.

There was a considerable gender bias towards females. This was not in itself surprising as there is a greater tendency for females to seek orthodontic treatment than males (Chestnutt et al., 2006). Elimination of the probands from the calculation produced an increase in the Female to Male ratio of the sample. There is no plausible explanation to date for this bias. However there was no significant difference between the risks to relatives of male or female probands. This reinforced the evidence for a major genetic component. Were the trait to be multifactorial, the risk would be higher for the relatives of the less susceptible gender (Pericak-Vance and Haines, 1998).

4.1.1.3 Segregation Analysis

The inheritance of EC in the families was modelled assuming a single major gene and with polygenic inheritance. The results of the segregation analysis implied that the sporadic model was rejected compared to all the genetic models. The dominant model fitted equally as well as the co-dominant model, suggesting that a dominant mode of inheritance adequately described the inheritance of EC in these pedigrees. The polygenic model was rejected, and there was no evidence for a polygenic component in addition to a dominantly inherited major gene.

4.1.1.4 Twin study

The inclusion of the first five sets of monozygotic and the first six sets of dizygotic twins was serendipitous. These families were selected on the basis of affected probands only. The low concordance shown by the monozygotic twins (Table 10) was indicative of epigenetic or other factors influencing the eruption of teeth and was at variance with the epidemiological evidence of a major genetic component to the condition. Furthermore, the incidence of monozygotic twinning for the Maltese Islands is 4.5 per 1000 (Savona-Ventura and Grech, 1988). The prevalence of monozygotic twins selected consecutively, (i.e. the first five sets) in this sample was nearly double that number and the difference was statistically significant, ($p = 0.01$)

Monozygotic twins share identical DNA sequences, however they are often discordant for certain phenotypes including complex diseases. Twin discordance has been reviewed by Silva (Silva et al., 2011) (Table 11).

Table 11. Causes for monozygotic twin discordance (Silva et al 2011).

Zygotic	Chromosomal mosaicism (pre and post-twinning)
	Unequal allocation of stem cells between twins
	Unequal placental territoriality
Epigenetic	Differential DNA methylation
	Skewed X inactivation
	Genomic imprinting
Mitochondrial	Unequal distribution of mitochondrial DNA
	Different level of heteroplasmic mutations
Genomic DNA	Post zygotic gene mutation
	Transposition of mobile elements
	Differential triplet repeat expansion
	Copy number variations

In the case of EC, the segregation analysis showed the disorder most likely to be genetic, autosomal dominant and unlikely to be due to a de novo mutation. The possible means to explain discordance in these twins may be heritable epigenetic mechanisms such as differential DNA methylation. Copy number variations (CNV) may also be a possible explanation (Notini et al., 2008). De novo, asymmetric CNV formation may be associated with the twinning event (Bruder et al., 2008). There was no pattern of anticipation, i.e. increasing severity or earlier incidence of the disease through generations and therefore no evidence from the pedigree analysis to support repeat expansion as a cause for EC.

Methylation of Cytosine-Guanidine (CG) dinucleotides found in stretches of multiple repeats (CpG islands) affects regulatory regions. Certain imprinting-associated diseases such as Beckwith-Wiedemann Syndrome are associated with multiple births and discordance in monozygotic twins (Elliott and Maher, 1994). Methylation aberrations may be transmitted through generations

(Rakyan et al., 2001, Sandovici et al., 2003). Incomplete erasure of methylation patterns in the germline may give phenotypes exhibiting variable expression and incomplete penetrance (Kearns et al., 2000). Incomplete penetrance is usually attributed to differences in action of quantitative trait loci however there is no evidence of polygenicity in this case. Promoter methylation is a possible explanation for the highly variable expression and incomplete penetrance seen in EC subjects.

4.1.2 Conclusions

The evidence gathered from analysis of the pedigrees supported the hypothesis of a genetic aetiology for EC, with a single locus being involved. The most likely mode of transmission was autosomal dominant. On a population level, there was a mutation frequency of 11%, with penetrance of 36% in mutation carriers. There was no indication of a major environmental component in this sample. Further investigation into the molecular aetiology of EC is therefore justified and parametric linkage analysis on an informative family is feasible.

4.2 Genome wide linkage analysis

4.2.1 Results

All quality assurance checks were completed successfully on pedigree V1. Pedigree V2 may have included a phenocopy and was only used to confirm the quality control results on V4. No analysis was carried out on Pedigree V3 as the separation of the two families would result in greater loss of information than would be gained by the inclusion of individuals 37, 49 and 410. The

analysis of V6 would not give any additional information over that provided by V7; this pedigree was also discarded.

4.2.1.1 LD quality control

The calculation run on V4 10 Kb gave a narrow peak of 60 Kb with a LOD of 3.3 on chromosome 1p21. Increasing the marker spacing to 50 Kb gave a smaller peak in a different region. The width of the peak and the shift in position[3] led to the suspicion that this peak may be an artefact caused by LD (Table 12). This was confirmed by a similar calculation using V2 and the LD pruned marker set, which gave no positive peaks at all.

Table 12. The change in size and position of the peak with a spaced marker set.

Pedigree	Start Position	End Position	Size (b)	LOD
V4 10 Kb	95,352,399	95,981,826	629,427	3.3
V4 50 Kb	85,332,120	85,553,405	221,285	2.5

4.2.1.2 Analysis of V8

Maximal LOD scores of 1.5 were seen on Chromosomes 4, 5, 6, 11, 13 and 20. (Table 13, Figure 7A).

[3]All chromosomal positions refer to Genome Reference Consortium build 37.

Table 13. Positive LOD scores Pedigree V8.

Pedigree	Chr	Start Position	End Position	Size (b)	LOD
V8	4	141,629,176	146,281,857	4,652,681	1.5
V8	5	178,621,576	180,692,321	2,070,745	1.5
V8	6	143,837,815	153,851,594	10,013,779	1.5
V8	10	15,264,688	15,292,006	27,318	0.9
V8	11	20,154,539	36,454,231	16,299,692	1.5
V8	13	107,344,212	115,106,996	7,762,784	1.5
V8	20	35,972,287	59,411,348	23,439,061	1.5

4.2.1.3 Analysis of V7

With the reduced penetrance dominant model, (AA 0.00; AB 0.90; BB 0.90) a LOD score of 2.6 on Chromosome 11p13-15 was seen, as well as a LOD of 2.5 on Chromosome 20q.13. (Table 14, Figure 7B). The V7 analysis supported that of V8 in that the peaks on Chromosomes 6 and 11 remained broad. The areas of greatest interest were therefore the peaks on Chromosomes 6 and 11. There was a narrow peak on Chromosome 20 which reached a LOD of 2.5 and therefore bore investigating too.

Table 14. Positive LOD scores Pedigree V7

Pedigree	Chr	StartPosition	EndPosition	Size (b)	LOD
V7	4	141,692,883	146,281,857	4,588,974	0.6
V7	5	178,621,576	180,692,321	2,070,745	1.6
V7	6	143,837,815	153,851,594	10,013,779	1.6
V7	11	20,020,564	36,454,231	16,433,667	2.6
V7	20	55,026,741	57,832,175	2,805,434	1.6
V7	20	57,832,175	57,941,413	109,238	2.5
V7	20	57,941,413	59,411,348	1,469,935	1.6

4.2.1.4 Haplotyping

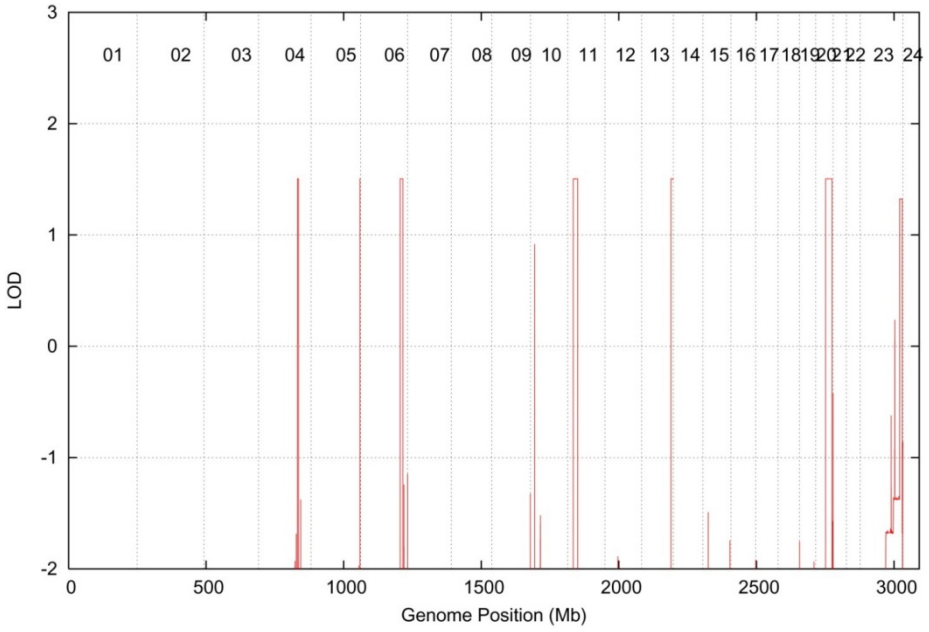
Haplotyping of the areas with the highest LOD scores in pedigrees V7 and V8 was carried out using all markers in order to identify the regions of interest. A common haplotype was seen in V7 on Chromosome 6 for all affected members plus #36 and 411, Chromosome 11 for all affected members plus individual #43 (Fig 9), plus another on Chromosome 20 for all affected individuals and #38 #41 #411 and #413 (Figure 10).

4.2.1.5 Exclusion of linkage

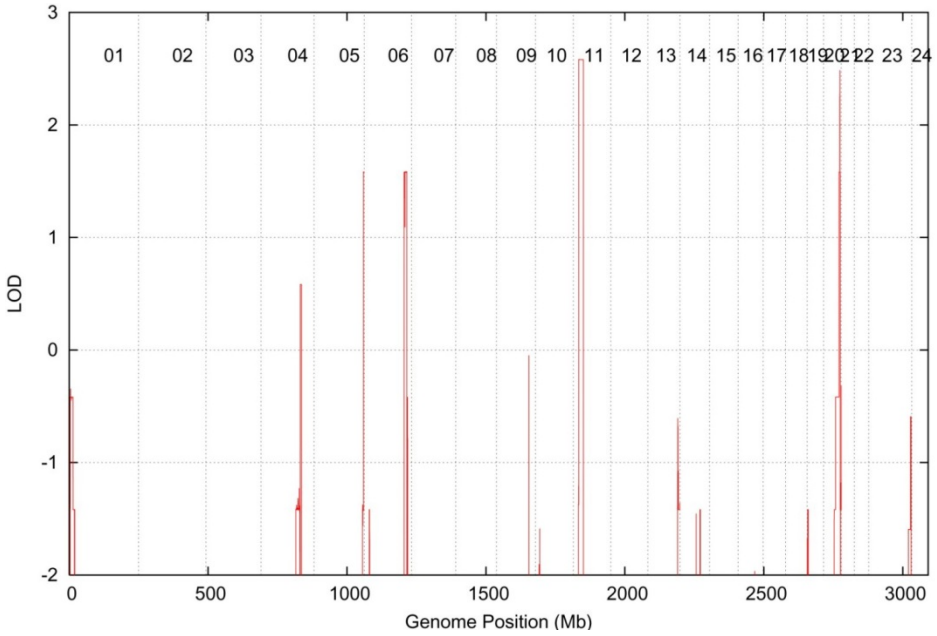
Using data from pedigree V8, the LOD scores on markers spanning the genes *SH3BP2, PTHR1, MSX1* and *PAX9* were found to be less than -2. This is the accepted score for exclusion of linkage (Morton, 1955) (Table 15).

Table 15. Linkage exclusion of eruption related genes.

Pedigree	Gene	1st Marker	LOD	2nd Marker	LOD
V8	*SH3BP2*	rs433030	-2.5	rs1263416	-2.5
V8	*PTHR1*	rs1531136	-2.5	rs2132173	-2.5
V8	*MSX1*	rs4348044	-2.5	rs3775261	-2.5
V8	*PAX9*	rs2295218	-6.5	rs17104928	-6.5



A



B

Figure 7. (A) V8 LOD plot. The maximum LOD from this pedigree was 1.5 and maximum scores were seen on chromosomes 4, 5, 6, 11, 13 and 20. (B) V7 LOD plot. The maximum LOD from this family was 3.9 and scores of 2.6 and 2.5 were seen on chromosomes 11 and 20 respectively.

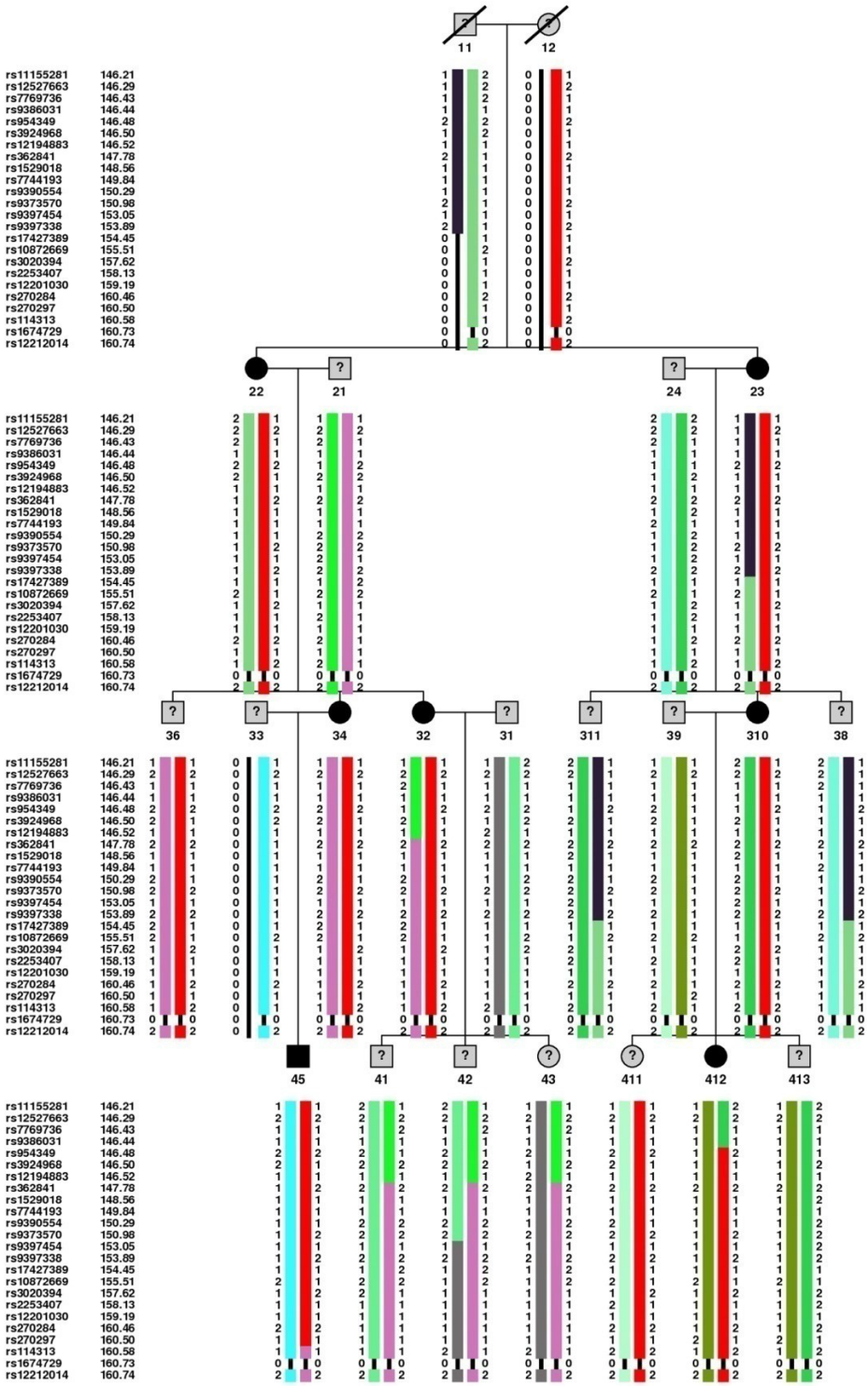


Figure 8. The haplotype on Chromosome 6. The disease haplotype is in red, inherited from the maternal founder. Markers have been selected for illustration, not all are shown.

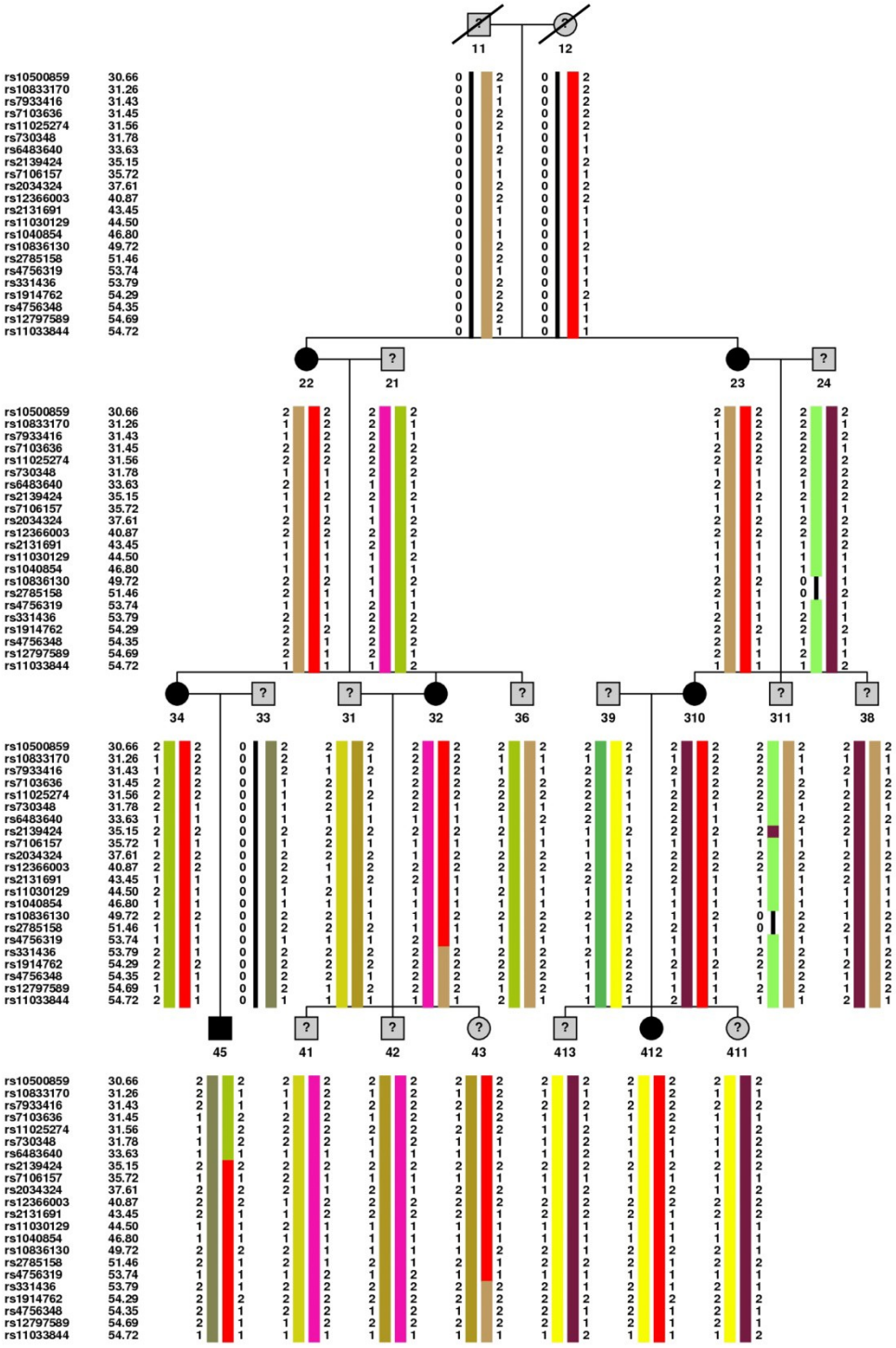


Figure 9. The haplotype on Chromosome 11. The disease haplotype is in red, inherited from the maternal founder. Markers have been selected for illustration, not all are shown.

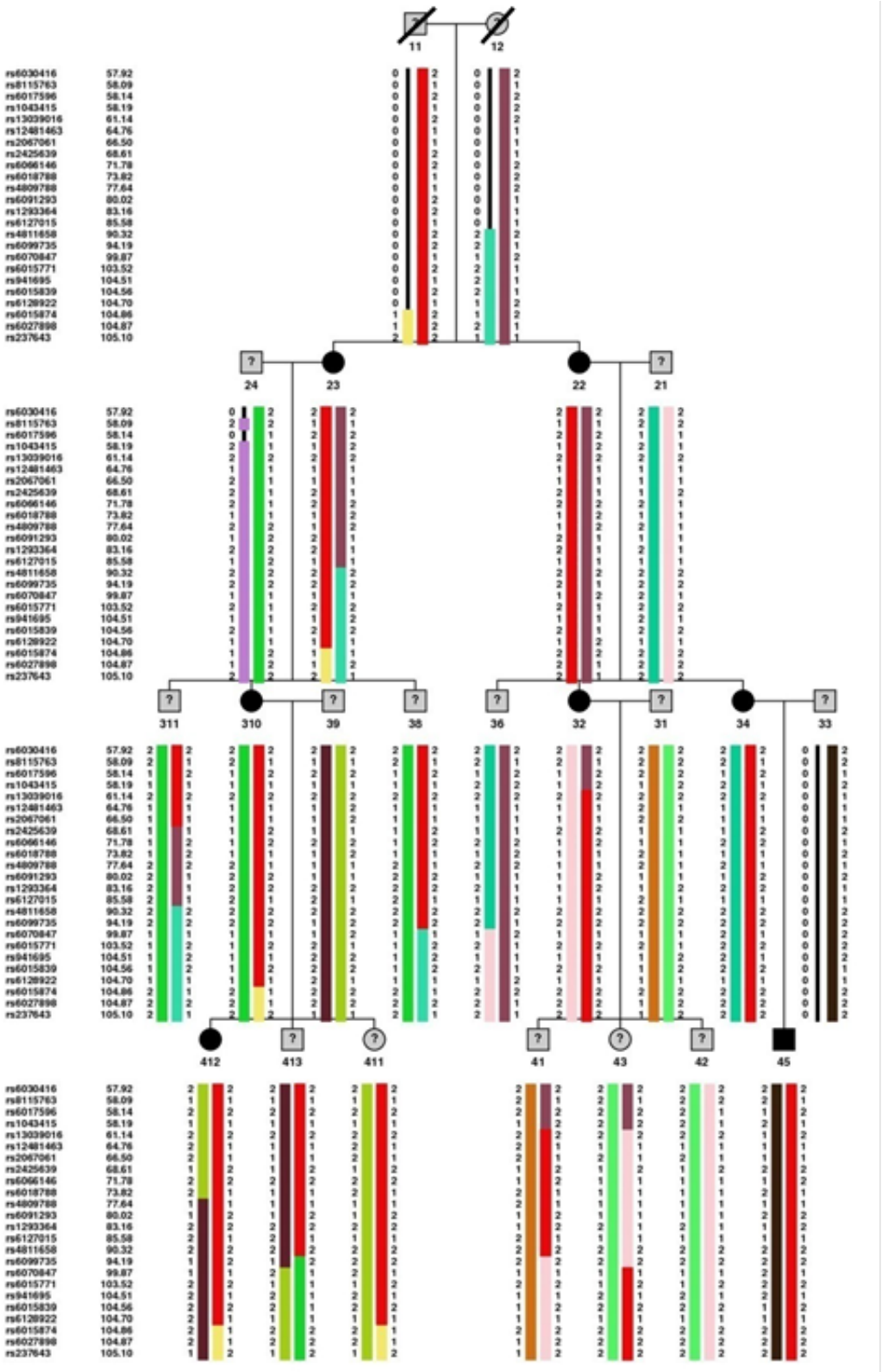


Figure 10. The haplotype on Chromosome 20. The disease haplotype is in red, inherited from the paternal founder. Markers have been selected for illustration, not all are shown.

SNPs flanking those indicated by Geller et.al. were also tested for exclusion. (Table 16).

Table 16. Linkage exclusion of SNPs flanking those identified by Geller et al.

Pedigree	Chr	SNP	1st Marker	LOD	2nd Marker	LOD
V8	1	rs2281845	rs12561765	-6.1	rs1059625	-6.1
V8	2	rs4491709	rs2888447	-6.1	rs7561690	-6.1
V8	10	rs7924176	rs1874152	-6.2	rs10509350	-6.2
V8	12	rs12424086	rs1042725	-9.6	rs10878353	-9.6

4.2.1.6 Copy Number Variations

No significant copy number events were evident in the areas defined by the linkage analysis.

4.2.2 Discussion

4.2.2.1 LD and quality control

The results of the quality control V4 analyses indicate that LD may be in operation even with an inter marker spacing of 50 Kb. The narrow initial peak obtained in V4 10k and the subsequent shift and further narrowing of this peak is indicative of this phenomenon. The use of a pruned marker set confirmed this result. In the V2 analysis, using the LD pruned marker set and an autosomal dominant model with reduced penetrance, the LOD score of 3.3 obtained in V4 10k dropped to -1.40. The criterion for exclusion of linkage is a LOD of -2, so the LOD score indicated an area of minimal importance. Comparing the two methods used to reduce the effects of LD, selection of marker SNPs on the basis of genetic distance is a more reliable method than selection by physical distance.

LD is a recognised problem in pedigrees where parental genotypes are missing. This is not the case here. The reason why such strong LD effects should be operating in an informative pedigree is puzzling. The Maltese population is young, a population bottleneck occurring just over 500 years ago. This equates with approximately 20 generations. LD blocks may be large, due to insufficient time elapsed to allow decay. LD has been shown to be present in isolated populations, however the amount of LD exhibited is variable and should only be in operation in the case of rare alleles (Kruglyak, 1999). The genotyping of more Maltese families may throw light on this problem.

4.2.2.2 Genetic model

Although segregation analysis indicated a single major gene this is not necessarily so. The trait is a complex one with reduced penetrance, variable expression and an environmental or epigenetic input. These factors may confound linkage analysis by incorrect identification of silent carriers as unaffected individuals. The mechanism of incomplete penetrance is complex and poorly understood and may be a feature of monogenic diseases, occurring despite relative genetic and environmental homogeneity (Pereira et al., 1994).

Parametric LOD score analysis is the most powerful method for establishing linkage; however the genetic model, i.e. the mode of inheritance, allele frequency and penetrance must be specified. Indeed it has been shown that specifying the wrong dominance mechanism of a linked locus decreases the expected LOD score however the wrong penetrance assumptions alone had relatively small effects on the magnitude of the *Z* value (Clerget-Darpoux et al.

1986). Multipoint analysis is more robust than two-point analysis to misspecification of allele frequencies and statistical fluctuations at individual markers. (Kruglyak et al., 1996). Furthermore, if a disease is caused by two epistatic loci or two independent loci, single-locus linkage analysis is nearly as powerful as two-locus linkage analysis in detecting linkage of one of the disease loci to a marker locus. However, this result assumes that the mode of transmission (dominant or recessive) of the linked locus is correctly specified (Goldin and Weeks, 1993). Complex traits may be analysed with parametric linkage analysis as long as the marker linked to the gene is inherited in a Mendelian fashion (Greenberg et al., 1998). Thus, linkage analysis has been shown to be robust to a number of incorrect specifications; however the mode of inheritance must be correct. The segregation analysis had identified a dominant mode as being the most likely, this supporting previously published results (Svinhufvud et al., 1988).

The analysis of V8 gave maximal LOD scores of 1.5 on chromosomes 5,6,11 and 13. Parametric affected-only analysis is the most conservative model but the chance of false positives is increased and the peaks are wider. The increase in LOD scores in the ‘affected only’ analysis is most likely due to the elimination of silent carriers.

Inclusion of unaffected members in the analysis will increase the accuracy of disease haplotype estimation in affected individuals. In a family such as this, silent carriers must be taken into account, allowing for reduced penetrance in the genetic model. On the other hand, over-relaxation of the penetrance vector may generate false positive results and indicate a ‘disease’ haplotype where

there is none. Indeed, the peaks on chromosomes 10 and 13 apparent in V8 disappear in V7, while suggestive LOD scores of 2.6 and 2.5 on chromosomes 11 and 20 are apparent in V7 but not in V8. Furthermore, these peaks are marginally narrower in V7 than in V8. The two analyses agree in that the peaks on Chromosomes 6 and 11 are both broad and very similar in both analyses. These therefore are the areas of greatest interest. The narrow LOD 2.5 peak on Chromosome 20 does not resemble a conventional LOD peak, however as precautions have been taken to remove LD, this area also bears investigation.

4.2.2.3 Phenotype

The phenotype of individual #41 is different from the rest of the affected members in the pedigree in that the maxillary canine had erupted high and buccal. Buccal canine eruption is generally associated with inadequate arch space (Chaushu et al., 2003a). However buccal ectopic canine eruption may occur as a distinct phenomenon, in spite of adequate space within the dental arch and is associated with an increased prevalence of lateral incisor anomalies (Chaushu et al., 2009). The high prevalence of associated dental anomalies in both conditions points to a common aetiology and indeed buccal EC is frequently found to occur in the same individual together with PDC. Crowding was not a factor in the position of this tooth, this being accommodated into the arch with the aid of an upper fixed appliance without any form of space creation, despite all teeth being of normal size. All criteria for inclusion were thus satisfied. Notwithstanding, it was considered unusual that a family exhibiting consistent phenotypic expression in seven other members should

display such a variation. The problem could have been dealt with by allowing for phenocopies in the genetic model, thus relaxing the parameters; however this may have masked potentially interesting areas by decreasing the LOD score. The alternative of eliminating #41 as an affected member seemed a better solution and indeed the analysis of pedigree V7 gave a LOD of 2.6 on a 16.5 megabase (Mb) region spanning Chromosome 11p13 to 11p15, a LOD of 2.5 on 20q13 and a LOD of 1.6 on Chromosome 6.

4.2.2.4 Results of haplotyping

Haplotyping showed all affected members plus #36 and 411 to carry a common haplotype on chromosome 6 (Figure 8), all affected members plus one unaffected individual, #43, to carry a common haplotype on chromosomes 11 (Figure 9) and all members plus #38 #41 #411 and #413 to carry a common haplotype only on chromosome 20 (Figure 10). The area of haplotype common to all affected members extended from marker rs11025381 to rs4471390 on chromosome 11. The haplotyping confirmed the presence of silent carriers, hindering interpretation of the result.

4.2.2.5 Identification of a candidate gene

In pedigree V7, the area covered by the positive LOD score on Chromosome 11 is 16 Mb long and contains 59 known genes; the area on Chromosome 20 is 4.4 Mb and contains 46 genes. Sequencing these genes would be time consuming and costly. Reduced penetrance and LD between the SNPs will not allow any further refinement of the fields by incorporation of more markers. It

may be possible to genotype these areas with alternative markers such as microsatellites, which are more informative than SNPs, though these are more widely spaced and more costly to process. Fine mapping is also an option but would require the genotyping of more individuals. With the advent of next generation sequencing, it would be more cost effective to employ exome sequencing in order to identify the disease gene.

4.2.2.6 Exclusion of linkage

The LOD scores of the SNPs flanking the genes *MSX1, PAX9, PTHR1* and *SH3BP2* were evaluated. No area was found to give a LOD of over -2 in pedigree V8. This is interpreted as exclusion of linkage (Morton, 1955). This result does not support the hypothesis of Peck et al, who postulated that the *Msx1* and *Pax9* genes may be responsible for PDC and transposed canines, both variants of EC (Peck et al., 2002). Furthermore, EC does not seem to share any genetic aetiology with Idiopathic Failure of Eruption or Paget's disease of bone. Similarly, the SNPs associated with dental maturation (Geller et al., 2011) are excluded from linkage in this family.

4.2.2.7 Copy number variations

No significant events were detected by the Copy Number probes. This is not surprising as, although CNVs contribute to human variation and play a part in the discordance of twins, there is no supporting clinical evidence. A major limitation of this analysis is that as it was carried out on a case-control basis; the numbers of both cases and controls would be far too small to be interpreted

reliably. Nevertheless, there is no evidence to suggest that CNVs may play a part in the aetiology of EC.

4.2.3 Conclusions

A number of areas of interest have been identified by a conservative 'affected only' analysis. These results are supported by linkage analysis of the extended family and areas on Chromosomes 4,6,11 and 20 bear further analysis. Genes associated with oligodontia have been excluded from linkage in this family as have SNPs associated with normal eruption. Nor is there evidence for any role of CNVs in the aetiology of the trait

4.3 Exome Sequencing

Linkage analysis has been successful in identifying many disease genes. However even under ideal circumstances the disease locus can rarely be mapped to an area of less than 1 Centimorgan (cM). One cM is a unit of genetic distance, not physical separation, however approximates in size to one Mb. An area of this magnitude may contain hundreds of genes and conventional Sanger sequencing would entail considerable time and expense. The emergence of next-generation massively parallel sequencing and high throughput sequence capture methods has made the targeted capture of all known exons highly cost effective. Protein coding regions constitute only 1% of the total human genome but are responsible for 85% of mutations (Choi et al., 2009). Mutations having major genetic effects generally disrupt protein coding regions or exon splice sites, mutations in regulatory regions tending to

have lesser effects (Ng et al., 2009). Exome sequencing enables identification of mutations in families that are insufficiently informative and not useful for linkage or other positional cloning methods. As only a small number of individuals need to be sequenced, it is also much quicker and cheaper (Singleton, 2011). Thus exome sequencing is a highly efficient method of capturing a large percentage of genomic variation and has reversed the recent decline in interest in monogenic, Mendelian disorders (Antonarakis and Beckmann, 2006).

4.3.1 Exome capture

Massively parallel sequencing has made possible the processing of enormous quantities of DNA templates within a much shorter timeframe and lower cost than Sanger chain termination technology. It has made possible the sequencing of whole genomes within a single centre, however the resources required for sequencing, storing and analysing the data from large numbers of genomes is beyond the reach of most centres. Target enrichment methods may be used to maximize the cost-efficiency of the process by selecting the portion of the genome most likely to yield information about a disease, that is, the exome.

Many different methods of target enrichment have been described (Mamanova et al., 2010) but the system that has been commercially applied to exome sequencing is the hybridisation method. Two variations exist, microarray-based and liquid based. In microarray-based capture, sonicated DNA is flowed over the probes, where complimentary sequences bind. The unbound DNA is washed away and the bound DNA amplified and sequenced. Solution-based

methods use biotinylated probes which bind to their complementary sequences on the DNA. Streptavidin coated beads are used to separate the hybridised probes and the DNA is eluted off. Both RNA and DNA probes are in commercial use. Liquid-based capture techniques are preferable in that they are less labour intensive, specialized hybridisation equipment is not required and the quantity of DNA required is considerably less.

4.3.2 Limitations of Exome Sequencing

As with all other techniques whole exome sequencing has its drawbacks (Bamshad et al., 2011). Our knowledge of all truly protein-coding exons in the genome is still incomplete, so current capture probes can only target exons that have been identified so far. There is a degree of variability between capture arrays, though commercial kits now target, at a minimum, all of the RefSeqcollection and an increasingly large number of hypothetical proteins. Considerable uncertainty remains regarding which sequences of the human genome are truly protein-coding and regulatory regions and untranslated regions are not sequenced.

The efficiency of capture probes varies considerably and some sequences fail to be targeted by capture probe design altogether. Not all templates are sequenced with equal efficiency and not all sequences can be aligned to the reference genome so as to allow base calling. On average, 82% of the genes on the array have at least 90% of bases called. Not all variants may be captured

It is difficult to detect small insertion/deletion variants (indels) with the short sequence reads generated by NGS technologies (Ng et al., 2010a). Exome

sequencing is unable to detect structural variants or chromosomal rearrangements (Ku et al., 2011) and fails to identify repeat mutations such as triplet repeats in spinocerebellar ataxia and copy number variants (Singleton, 2011). AT and GC rich regions are not easily sequenced (Bloch-Zupan et al., 2011).

As the number of sequenced genomes increases, so does the possibility of incorporation of pathogenic alleles at appreciable frequency. This particularly applies to recessive disorders (Bamshad et al., 2011). Finally, genetic heterogeneity may confound results (Ng et al., 2010a).

4.3.3 Filtering

A large amount of data is generated which must be filtered to narrow the search down to a manageable number of candidate genes.

The polymorphisms are filtered first against a control set, obtained either from public databases, assuming that there are no disease alleles present, or from a control cohort of individuals known to be unaffected. It may be assumed that common variants in the population are not likely to be the cause of rare Mendelian diseases and may be filtered out. This leaves the 'novel' polymorphisms and rare variants as potential candidates. Nonsynonymous variants, splice acceptor and donor site mutations and coding indels may be prioritised, on the assumption that synonymous variants are far less likely to be harmful.

Further filtering may be applied by comparing of affected individuals, either on a whole exome basis if unrelated or by comparison of areas IBD in family-

based studies. As affected family members may be assumed to carry the same rare variant, family studies require smaller sample sizes than studies of unrelated individuals (Wilson and Ziegler, 2011).

In families where multiple individuals are affected with a common trait, the most distally related individuals are sequenced, presuming that the more distantly related the individuals, the fewer genetic variants they share. However, even distantly related individuals may share many variants and will require more filtering. Homozygosity mapping may be used in the study of recessive disorders, on the assumption that the mutation must be present in both alleles.

An alternative, family-based approach, which is used to identify *de novo* variants, involves sequencing parent–offspring trios in which only the offspring is affected. This strategy has been used to identify candidate genes for several complex traits (Vissers et al., 2010)

Where unrelated individuals are studied, selection of phenotypes at the extreme ends of the distribution curve will make it more likely that they share the same rare variant. Careful selection and ranking of phenotypes will avoid the problem of genetic heterogeneity (Ng et al., 2010a). Case-control designs may also be used and as the cost of sequencing continues to fall and public databases increase in size, may become more popular.

Sequence variations are reported in public databases such as dbSNP or the 1000 genomes database. The relevance of the variation to the disease may be gauged by expression data or functional stratification and prediction of the functional consequences may be carried out using programs such as SIFT

(Kumar et al., 2009) or PolyPhen2 (Adzhubei et al., 2010). Both these programs are based on the degree of conservation of amino acid residues in the sequence alignments. PolyPhen2 also takes into account CpG context and predicts the alteration in structural features that may affect protein function. This function is naturally dependent on knowledge of the structural features of that protein. It should be noted that computational algorithms have high rates of false positive and false-negative predictions (Wei et al., 2010, Mathe et al., 2006). Although it is difficult to give an exact numerical value, it is likely that the false-negative and false-positive rates are at least 20% for whole exome sequencing (WES) data (Robinson et al., 2011).

4.3.4 Results

The percentage of the exome covered 10x was 96.55, that covered 20x was 93.96. The mean coverage was 215x. The numbers and types of variants identified in both individuals are given in Table 17. Three heterozygous novel variants were found to be common to the two affecteds (Table 18) as well as twenty-five other variants with a frequency of 0.01% or less (Table 23, in Appendix 2). None of these genes were located within the linkage areas in the analysis of Pedigree V8. The LOD score for the variants in Chromosomes 6 and 22 in Pedigree V8 was -2.5, indicating exclusion of linkage to the trait in the affected members of this family. Four variants common to both members were found in the areas identified by linkage analysis (Table 19). One variant was found on the *PAX9* gene. This variant is a G-C transition present on the dbSNP database as rs4904210, with a population frequency of 0.3428.

The four genes in the linkage areas were prioritised, in order of frequency and relevance to function. None of the variants were predicted to be functionally relevant (Table 20).

4.3.4.1 Sanger Sequencing results.

The variants on *ANO5, PPP1R14C* and *EPM2A* were found to be present in all affected members. The SNV on *VSTM2L,* did not segregate with the affected members. Combinations of these variants were also present in first degree relatives of affected members, but there was no unaffected member with both variants of *ANO5* or *PPP1R14C* (Figure 11). During the course of the investigation, member #47 developed a macrodont upper right central incisor, however none of the variants were found in this individual.

The variants on *PPP1R14C, ANO5, EPM2A* or *VSTM2L* were not present in any of the unrelated individuals. However one novel SNV and one SNP were found in seven and eleven individuals respectively in exon 1 of *PPP1R14C* (Table 21). These are both missense variations predicted to be harmless. No variants were detected in the intron or the intron-exon boundaries of exon 17 of ANO5 in any of the eighteen unrelated individuals.

Table 17. Summary statistics for exome sequencing, giving the variants in each individual as well as those common to both.

Individual	**32**		**412**		**Shared**	
Variant	**Known**	**Novel**	**Known**	**Novel**	**Known**	**Novel**
All variants	24707	113	24298	316	15690	3
Heterozygous	15270	111	15165	311	7745	3
Homozygous	9437	2	9133	5	7945	0
Non frameshift	226	13	246	11	147	0
Heterozygous	129	12	161	10	72	0
Homozygous	97	1	85	1	75	0
Frameshift Indels	172	0	175	4	133	0
Heterozygous	57	0	54	3	27	0
Homozygous	115	0	121	1	106	0
SNVs	21478		21232	269	13599	3
Heterozygous	13335	81	13246	266	6742	3
Homozygous	8143	1	7986	3	6854	0
Synonymous	11098	31	10955	97	7060	1
Heterozygous	6893	30	6849	95	3526	1
Homozygous	4205	1	4106	2	3533	0
Non synonymous	10274	49	9966	164	6537	2
Heterozygous	6383	49	6255	164	3216	2
Homozygous	3891	0	3711	0	3321	0
Splice Site(10bp)	2791	18	2790	36	1792	2
Heterozygous	1723	18	1734	35	892	2
Homozygous	1068	0	1056	1	900	0
Unknown	40	0	30	0	20	0
Heterozygous	26	0	24	0	10	0
Homozygous	14	0	6	0	10	0

Table 18. Nonsynonymous novel variants common to the two affected members. The LOD scores of both V7 and V8 analyses indicate the areas are unlikely to be shared IBD.

Chromo-some	Position	Ref Allele	Alt Allele	Zygosity	Exonic/ Splicing	Details	LOD V8	LOD V7
Chr22	33,402,797	C	T	HET	splicing	SYN3 splicing	-2.495	-6.417
Chr22	38,379,730	C	A	HET	exonic	SOX10: nonsynonymous SNV NM_006941:exon2:c.G62T:p.R21L,	-2.495	-6.418
Chr6	13,711,706	T	G	HET	exonic	RANBP9: nonsynonymous SNV NM_005493:exon1:c.A32C:p.Q11P,	-2.495	-3.418

Table 19. The rare variants that fell within areas identified by linkage analysis. The variant on Chromosome 20 has a positive LOD score in the conservative V8 analysis but has a negative score in the V7 analysis.

Chromo-some	Position	Ref Allele	Alt Allele	Zygosity	Exonic/ Splicing	Details	LOD V8	LOD V7
Chr 6	146,007,358	T	C	HET	Exonic	EPM2A: Nonsynonymous SNV NM_005670:exon2:c.A376G:p.I126V, NM_001018041:exon2:c.A376G:p.I126V	1.505	1.581
Chr6	150,464,476	G	T	HET	Exonic	PPP1R14C: nonsynonymous SNV NM_030949:exon1:c.G148T:p.V50	1.505	1.582
Chr 11	22,284,483	T	C	HET	Splicing	ANO5: splicing NM_001142649:exon17:c.1798-9T>C, NM_213599:exon17:c.1801-9T>C	1.505	2.581
Chr 20	36,560,190	C	T	HET	Exonic	VSTM2L: Nonsynonymous SNV NM_080607:exon2:c.C275T:p.A92V,	1.505	-2.418

Table 20. The final list of candidate genes. A SIFT score of <0.05 predicts the mutation to be deleterious. PolyPhen scores range from 0 (benign) to 1 (deleterious).

Gene	Frequency (1000 Genomes/Db SNP)	Pooled frequency in 50 Maltese individuals	LOD Score V7	Function	SIFT Score	PolyPhen2 score
PPP1R14C	0.004	0.01	1.6	Protein phosphatase inhibitor	0.45	0.012
ANO5	0.0005	0.0	2.6	Development of musculoskeletal system	n/a but within canonical splice site	
VSTM2L	0.0045	0.01	1.6	Potential modulator of NFKB signalling, involved in cell proliferation.	0.07	0.355
EPM2A	0.0005	0.0	1.6	Involved in neuronal response to misfolded proteins	1	0

Table 21. The SNVs in PPP1R14C exon 1 in eighteen unrelated EC cases (P1-P18). One SNV is novel and rs2297672 has a MAF of 0.39. N = variant absent, Y = variant present.

DbSNP ID	Novel	rs2297672
Position	6:150464335_G/A	6:150464356; G/A
P1_1	N	**Y**
P1_2	N	**Y**
P1_3	**Y**	N
P1_4	**Y**	**Y**
P1_5	**Y**	**Y**
P1_6	N	**Y**
P1_7	**Y**	**Y**
P1_8	N	**Y**
P1_9	N	N
P1_10	**Y**	**Y**
P1_11	N	N
P1_12	**Y**	**Y**
P1_13	**Y**	N
P1_14	N	N
P1_15	N	**Y**
P1_16	N	N
P1_17	N	N
P1_18	N	**Y**

4.3.5 Discussion

There are three main exon capture kits on the market. They differ markedly in type of bait used to capture reads and also in density and target regions. Of these the Agilent Sure Select Human Exome Kit gives good coverage of Ensembl coding regions. Moreover the long RNA baits give it an edge over its competitors in detecting indels.

4.3.5.1 Quality control

Reliable data requires a depth of at least 5 reads per variant with a minimum of 20% of reads showing that variant for heterozygous SNVs and 80% for homozygous. The sensitivity to detect heterozygous variants with 10 reads is 78.6%, but increases to 95.2% at 20x and approximately 100% at 30x and greater (Choi et al., 2009). As 93.96% of the exome was covered 20X and 96.55% covered at least 10X, the data extracted was considered to be of good quality.

4.3.5.2 Filtering

Disease gene identification strategies for various models have been reviewed by Gilissen (Gilissen et al., 2012). The advent of next generation sequencing has shifted the problem from that of struggling to identify a candidate gene to that of elimination of sufficient numbers in order to arrive at a manageable quantity of candidates. Following quality control procedures, there remain approximately 25,000 variants in the exome and splice sites in each individual. The removal of synonymous variants reduces the number to 10,000. Further filtering involves the elimination of SNVs not common to both affected individuals. However as these are related individuals, large parts of the genome will be shared and several potential candidates may remain.

Linkage data may be used to identify areas which are identical by descent to all affected members. These areas may be prioritised and SNVs that lie outside the linkage peaks may be discarded. This strategy has been successfully used to identify a number of genes, including variants responsible for a tooth eruption

disorder (Yamaguchi et al., 2011). SNPs captured by exome sequencing may be used for linkage analysis, greatly reducing the search area. However as the markers in genotypes inferred from WES are sparser and less heterozygous than those from microarrays, linkage data from the latter is considered superior (Smith et al., 2011). Even within these linkage areas, large numbers of genes may remain. As the disease segregates within the family under study, it may be assumed that the same variant is responsible for the phenotype and variants not common to both exomes may be discarded.

Homozygosity or heterozygosity may be deduced from the mode of transmission and nonsynonymous variants are usually filtered out; bearing in mind that synonymous variants may cause alternate splicing, affecting protein properties. The assumption of a Mendelian gene with a major effect allows the field to be further narrowed to novel or rare SNVs. The frequency of identified variants is obtained from dbSNP or the 1000 genomes database. Caution must be exercised as many disease studies submit their findings to dbSNP. Similarly, samples submitted to the 1000 Genomes Project mostly have no associated medical or phenotype data. Thus both data bases may be ‘contaminated’ with causal variants where the quoted population frequency may not match that under study. Care must therefore be taken not to discard the true variant and search parameters may be relaxed to a frequency of <1% in the database.

Three novel variants as well as 25 other variants with a frequency of less than 1% were found to be common to the two exomes. These were not all in the linkage peaks, indeed, the LOD scores for the novel variants was less than -2.5 in the conservative ‘affected only’ V8 analysis and lower in the full family V7

analysis. Only four variants were inherited IBD in common to both genomes in analysis V8 and indeed the variant on Chromosome 20 has a negative LOD score in analysis V7. However the presence of novel variants common to 5th degree relatives may be significant and the segregation of these genes with the phenotype requires examination. The variant rs4904210 on *PAX9* is a common variant present in 34% of the population. It has been implicated as a risk factor for tooth agnesis (Wang, 2011). The V8 'affected only' LOD score indicates that it may not be common to all affected members, however the V7 'full family' analysis assigns a score of -0.5. This score is low, but is not exclusive of linkage.

4.3.5.3 Candidate genes

After applying the various filtering strategies, the final list of candidate genes totals four:

A low-frequency SNV in *ANO5*. This variation is present in the 3' splice region of exon 17. It is present at a very low frequency in the 1000 Genomes database. but is not present in dbSNP or in the pool of Maltese individuals. *ANO5* affects bone formation, particularly in the jaws. Mutations in exon 11 of *ANO5* are responsible for gnathodiaphyseal dysplasia, an autosomal dominant disorder affecting long bones and the jaws, in particular the formation of fibrous lesions in tooth-bearing regions. *ANO5* is expressed in bone and periodontal ligament, the latter being intimately involved in tooth eruption (Tsutsumi et al., 2004). It is also linked to the *NELL1* gene, which plays a key role in both intramembranous and endochondral ossification during early

development (Desai et al., 2006), its function and range of expression making it a good candidate. *ANO5* is in the region of the highest LOD score (2.58) in Pedigree V7 and the SNV did segregate with the affected phenotype, with only one unaffected family member, #43, carrying the SNV (Figure 11). No variations were found in the exon or intron-exon boundaries of exon 17.

A low frequency SNV identified in exon 1 of *PPP1R14C*, which falls in the region on Chromosome 6 that gave a LOD score of 1.6 in Pedigree V7, indicating an area of interest. This SNV is present at a frequency of 1% in a pool of 50 Maltese individuals. It is a protein kinase C-potentiated inhibitory protein for type 1 Serine/Threonine phosphatases, which works via the *MEK-ERK* pathway (Wenzel et al., 2007). It inhibits *PP1*, which acts on *AXIN*, which participates in both the *MAPK* and *WNT* pathways, with inhibition of *PP1* enhancing B-CATENIN degradation (Voronkov and Krauss, 2013). It also interacts with *SAP25*, which is involved in transcriptional repression mediated by *mSIN3A*, which in turn influences RUNX2 expression (Shiio et al., 2006, Westendorf, 2006, Wenzel et al., 2007). It has been found to be expressed in many tissues, including bone marrow and osteoblasts. Its known function and range of expression make it a viable candidate gene. It is not within the highest range of LOD score; however the variable penetrance of the condition makes silent carriers very likely and indeed the SNV did segregate with the affected phenotype, with two unaffected family members, #36 and #411, carrying the SNV. The scores given by both SIFT and PolyPhen2 do not indicate that protein function should be affected; however as previously stated, these predictions operate with a wide margin of error. Exclusion based on

prediction has proved to be erroneous (Ng et al., 2010b). Other SNVs were found in exon 1 of *PPP1R14C* in other affected, unrelated individuals, implying that other causative mutations may be present.

A low-frequency SNV in *VSTM2L*, present at 1% in the pool of Maltese individuals. Little is known about the function of this gene; however it may play a role in cell proliferation, in the modulation of NFKB function and in neuronal differentiation. It may also act to modulate the action of neuroprotective peptides (Mortazavi et al., 2006, Halsey et al., 2007, Rossini et al., 2011). The function of the gene is appropriate in that it may affect cell proliferation and that NFKB plays a part in osteoclastogenesis, which in turn affects tooth eruption. Both the SIFT and PolyPhen2 scores indicate low pathogenicity. However the SNV does not segregate with the phenotype, being present in five other unaffected family members and a founder (#39) (Figure 11). Nor were any mutations found in exon 2 or the intron-exon boundaries in the unrelated affected individuals. This is not altogether a surprising finding, as the V7 linkage analysis takes into account information from unaffected relatives in computing the haplotypes.

A low-frequency SNV in *EPM2A*, not present in the pool of Maltese individuals, which encodes LAFORIN, a dual-specificity protein phosphatase that hydrolyses phosphotyrosine and phosphoserine/threonine substrates. LAFORIN also binds complex carbohydrates and plays a role in glycogen metabolism. LAFORIN acts together with MALIN to accelerate the degradation of misfolded proteins and may be part of the neuronal defence

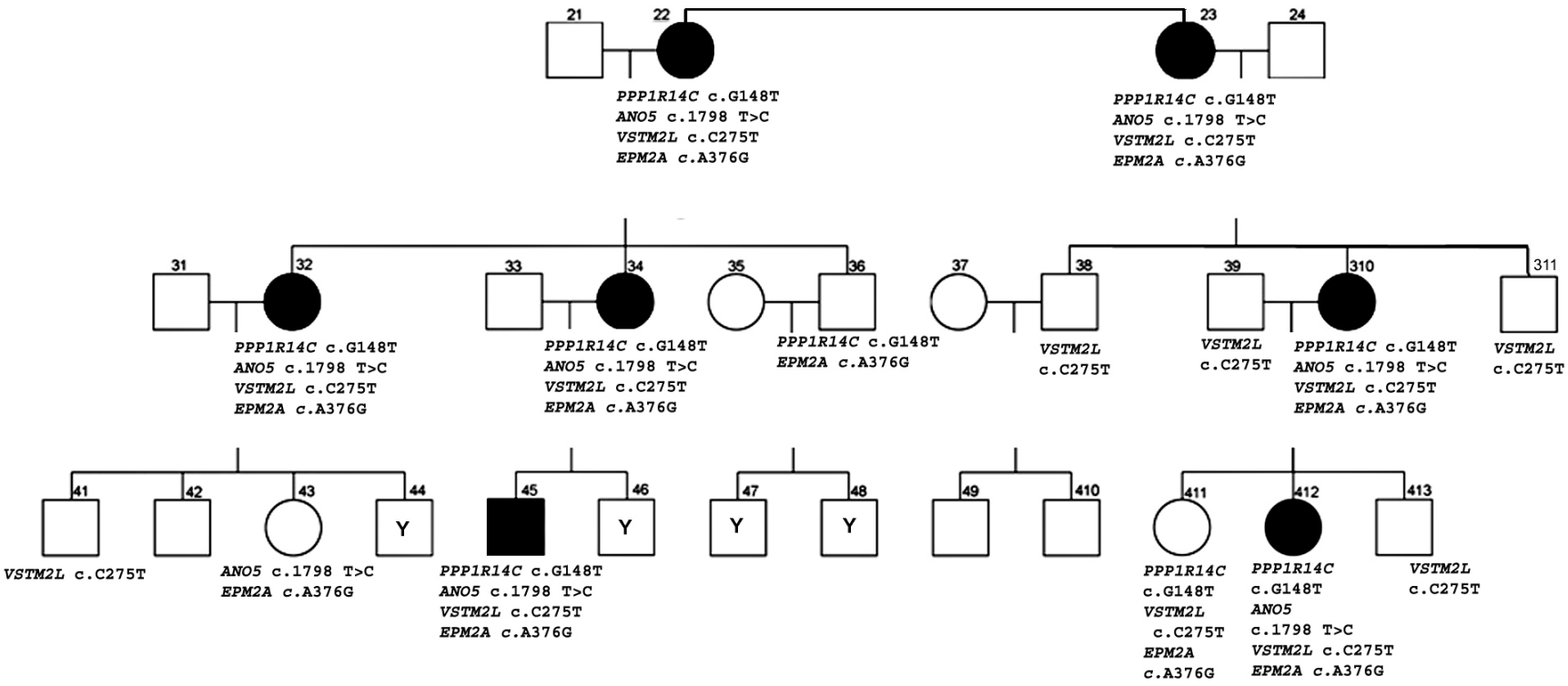


Figure 11. The distribution of the variants in the family. The SNVs in *ANO5* and *PPP1R14C* together are necessary to produce the phenotype.

mechanism against cytotoxic proteins (Garyali et al., 2009). However LAFORIN is not associated with bone or craniofacial development and has not been shown to be expressed in teeth, bone or the periodontal ligament. The known expression pattern and gene function is therefore not appropriate. The SNV on *EPM2A* did segregate with the phenotype however as its position is physically very close to *PPP1R14C*, this is not altogether surprising. No mutations were found in exon 2 or the intron-exon boundaries in the unrelated affected individuals.

The variant in *PPP1R14C* is present as a rare allele in a sample of the Maltese population; however neither of the SNVs in *PPP1R14C* or *ANO5* are present in a sample of 18 consecutive, unrelated sporadic EC cases, indicating that these variants will only account for a small percentage of the heredity of EC. In order to determine whether these two genes are expressed in tooth tissue or supporting structures, in situ hybridisation was carried out in human embryos. In order to confirm whether *PPP1R14C* or *ANO5* were expressed in teeth or supporting tissues, is situ hybridisation was carried out on both genes.

4.4 In situ hybridisation

The technique was published in 1969 and the use of radiolabelled RNA probes described in 1984 (Gall and Pardue, 1969, Cox et al., 1984). ISH is designed to localise gene expression in tissues. Specific probes are manufactured, which may be DNA or RNA, complementary to the sequence targeted. Double stranded DNA probes are prepared by cleaving off the DNA plasmid with restriction enzymes, single stranded probes may be manufactured by PCR

using Taq polymerase and antisense primer extension. RNA probes are prepared by cloning the cDNA template into a transcription vector flanked by two RNA polymerase sites. This enables both sense and antisense probes to be manufactured. DNA probes are easier to synthesise and more stable, however RNA probes are more specific.

These probes are labelled either radioisotopically and visualised by exposure to radiosensitve film, or with a nonradioactive substance such as digoxigenin, which may then be visualised by immunofluorescent staining. Isotope ISH is the more sensitive of the two techniques, and has the added advantage that the results may be quantified. Several isotopes may be used, however S^{35} gives the most rapid and sensitive results (Wilcox, 1993). The disadvantage of isotopic labelling is the use of biohazardous material and the limited shelf life of the probes. Nonisotopic probes are more stable, more sensitive and give much more rapid results, making them useful in clinical laboratories; however the results can not be quantified. Acetylation of the tissue will prevent backgound due to nonspecific binding and the addition of dithiothreitol to the hybridisation solution of S^{35} probes will prevent oxidization of the sulphur.

4.4.1 Results

Microphotographs of the in situ hybridisation may be seen in Figures 12 and 13. The background binding of both probes was similar, therefore non-specific binding was judged not to affect the result.

Thickening Stage

PPP1R14C expression is observed in both epithelium and mesenchyme.

Faint expression of *ANO5* is observed in epithelium.

Bud Stage

Strong expression of *PPP1R14C* is found in mesenchyme, while *PPP1R14C* shows weak expression in bud epithelium.

No expression of *ANO5* could be detected in tooth germs.

Cap Stage

PPP1R14C is strongly expressed in mesenchyme, whereas it shows weak expression in cap epithelium.

ANO5 is barely expressed in lingual cap epithelium.

Bell Stage

PPP1R14C is strongly expressed in ameloblasts and odontoblasts, whereas it shows weak expression in dental papilla cells and the dental follicle.

ANO5 is weakly expressed in ameloblasts and rostral dental papillae including odontoblasts. It is also weakly expressed in the dental follicle.

4.4.2 Discussion

Both genes are expressed in tooth tissue. *PPP1R14C* is expressed to varying degrees at all stages in epithelium, mesenchyme and dental follicle. *PPP1R14C* inhibits the protein phosphorylator *PPP1CA*, which plays a part in the *MEK/ERK* and WNT pathways. These pathways play a vital role in most developmental processes including bone and teeth, the strong expression of the gene here supporting *PPP1R14C* as a candidate. *ANO5* shows weaker expression, chiefly in epithelium and then in ameloblasts and odontoblasts at

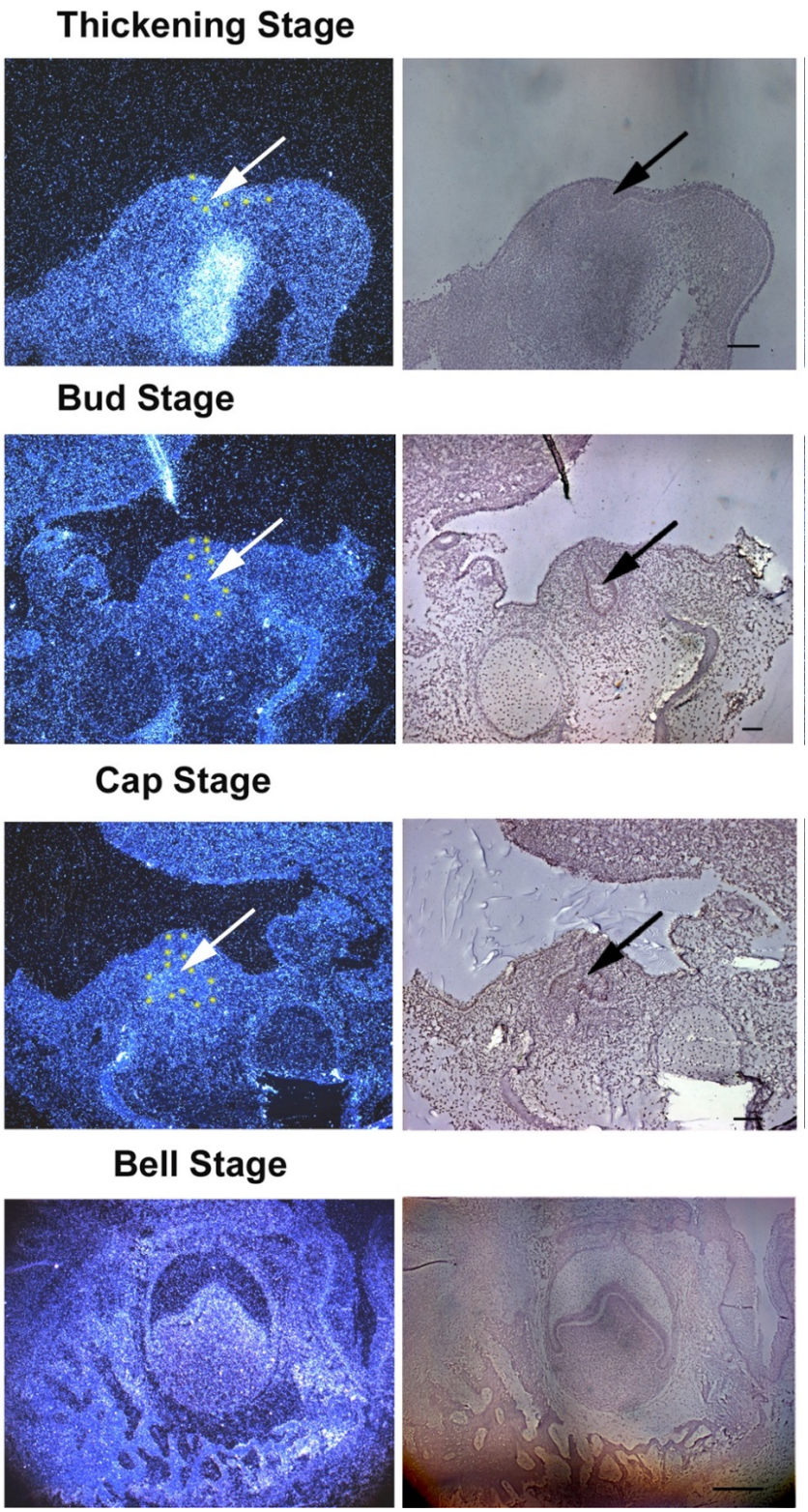


Figure 12. *ANO5* in situ. There is weak or no expression in dental tissue in the initial stages of tooth formation, however there is diffuse, weak expression in both tooth and dental follicle in the later stages of tooth formation. The yellow dots outline the developing tooth germ indicated by the arrows. Scale 200 µM.

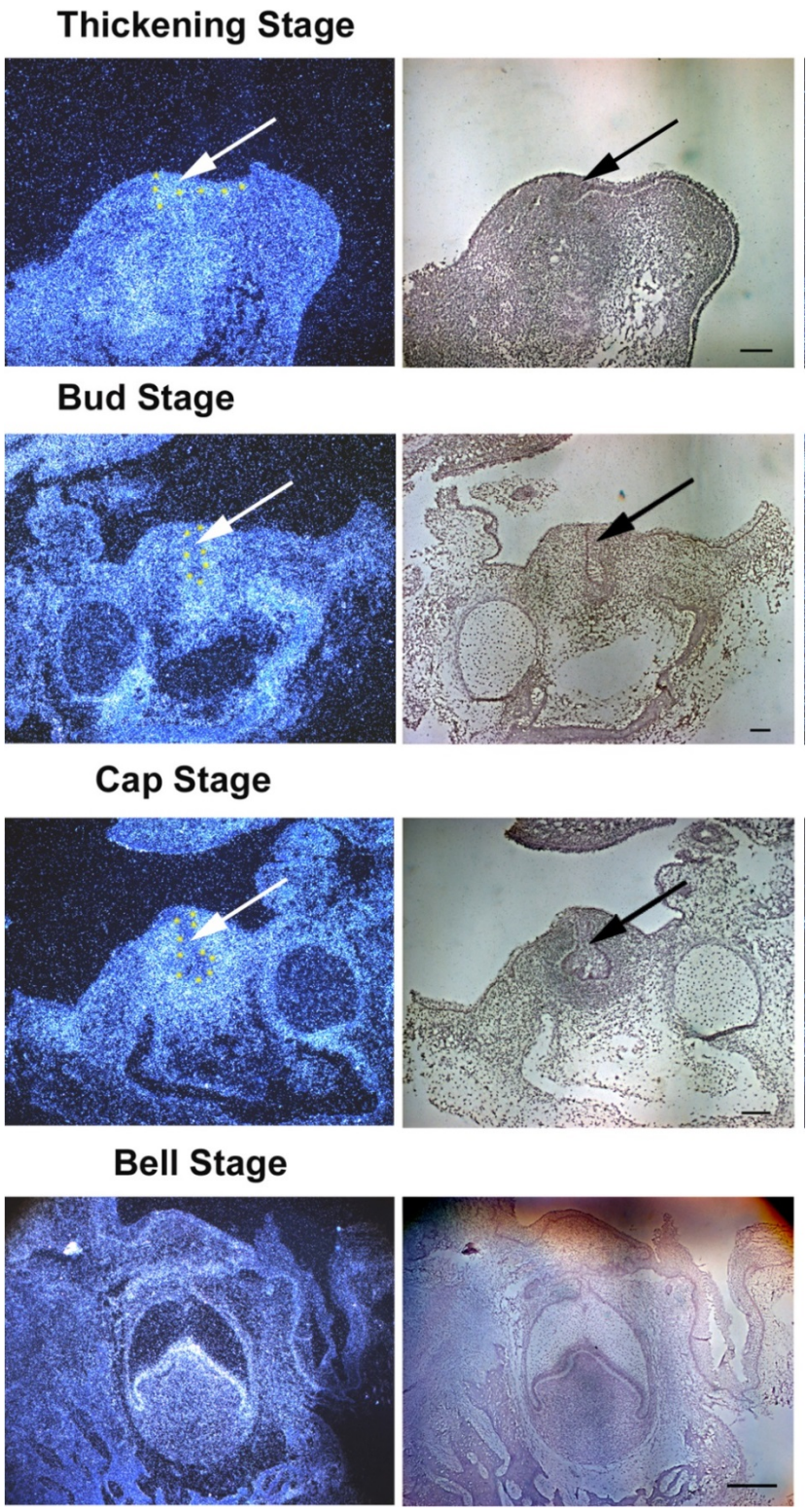


Figure 13. *PPP1R14C* in situ. There is strong expression in dental tissues throughout tooth formation. There is particularly strong expression in ameloblasts and weak expression in the follicle and supporting tissues at the bell stage of formation. The yellow dots outline the developing tooth germ indicated by the arrows. Scale 200µM.

the bell stage. Its expression in calcifying tissues lends support to the theory that this gene might function as an intracellular calcium-release channel and that it might regulate a calcium-dependent signaling pathway by modifying the intracellular calcium concentration (Tsutsumi et al., 2004). Its expression in dental follicle tissue is consistent with previous publications of expression inperiodontal ligament cells. As it is the dental follicle that controls tooth eruption, disruption of *PPP1R14C* or *ANO5* function could well be a cause of ectopic eruption. Either or both genes may be the cause of EC.

Sequencing the variations in the family shows that only the affected members have SNVs in both *PPP1R14C* and *ANO5* (Figure 11). Members with only one or the other variation are unaffected, as are those with combinations of the other candidate genes. It is apparent in Figure 11 that mutations in two genes, *ANO5* and *PPP1R14C* are necessary to precipitate the EC phenotype in this family, either mutation alone not being sufficient. Individuals #36 and #411 are both unaffected but carry the mutant form of *PPP1R14C*. Similarly, #43 carries the mutant form of *ANO5* but is also unaffected. Whether one gene has greater influence that the other is not clear, either could contribute in varying degrees to the phenotype. It is possible that one variation is insufficient, both being necessary for ectopic eruption to occur. This combination has been termed “pathogenic” digenic inheritance (Ming and Muenke, 2002).

Disorders with significant familial aggregation and which exhibit a strong genetic component may be diagnosed as monogenic. However these may exhibit complex inheritance patterns which depart from classical Mendelian inheritance. Thus the phenotypes of single gene disorders may manifest as complex traits (Dipple and McCabe, 2000) The causes for atypical patterns of

inheritance are several (Van Heyningen and Yeyati, 2004) however in this case a digenic form of inheritance may be in operation, with mutations in both genes necessary to produce the EC phenotype.

Tooth formation and eruption involve several complex pathways and these genes could act in the same or in related pathways to affect protein production. The involvement of more than one gene is consistent with the variable penetrance and expressivity of the trait, possibly complicated by the presence of modifier genes, which may affect the clinical manifestation of the disease altering the penetrance and/or expressivity of the phenotype (Gropman and Adams, 2007). #47 does not carry any of the SNVs identified, however may have related, but as yet unidentified causative variants.

These mutations have not been found to be present in several unrelated EC cases. If the condition were due to genetic drift with a single founder effect, then the same mutations would have been detected. It is likely that, due to the complex pathways leading to bone and tooth development, genetic heterogeneity exists, possibly due to multiple founders. Similar instances have been described for rare diseases in isolated populations (Richard et al., 1995).

4.5 Conclusion

The aims of the study have been reached, in that the evidence for a genetic aetiology for EC has been established by the familial relative risk and the segregation analysis, which has also determined the genetic model for the trait. Linkage analysis based on the findings of the segregation analysis has identified areas of interest in an extended Maltese family. The second aim has also been reached in that whole exome sequencing has implicated variants in

the genes *ANO5* and *PPP1R14C* in the aetiology of EC in this family. A digenic or oligogenic aetiology is in keeping with the clinical findings of variable penetrance and variable expression of EC However the discordance of identical twins points to a significant environmental or epigenetic component. The genes *MSX1* and *PAX9* have been excluded as a cause of EC, as have the SNPs identified by Geller (Geller et al, 2011) associated with normal eruption.

4.5.1 Future investigations

The presence of novel variants common to 5th degree relatives may be significant and the segregation of these genes with the phenotype requires examination The variant rs4904210 may be associated with hypodontia in this family and also bears investigation.

The discordance of the monozygotic twins is not consistent with a genetic aetiology. This phenomenon may either be due to the twinning process itself or may be an indication of environmental or epigenetic factors. The number of multiple births in the sample raises the possibility of epigenetic effects related to the twinning process or CNVs and methylome analysis as well as CNV analysis should also be undertaken in discordant twins.

The exons and intron-exon boundaries of both genes *PPP1R14C* and *ANO5* may be investigated in unrelated cases, as variants in other parts of the gene may be responsible. Other genes in the pathways of action of these genes may also be investigated. Custom targeted sequencing may offer a solution here.

Other extended families affected with EC may be recruited and investigations carried out in the light of the results of this study.

The products of the genes *ANO5* and *PPP1R14C* or any other genes identified in other affected individuals may be tested to see whether the RNA transcripts are altered. Quantitative Reverse Transcription Polymerase Chain Reaction (RT-qPCR) may be used to detect changes in RNA transcripts and their levels and determine whether both genes are at fault.

Chapter 5 References

Abecasis G R, Cherny S S, Cookson W O, Cardon L R. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002; 30(1): 97-101.

Abecasis G R, Wigginton J E. Handling marker-marker linkage disequilibrium: pedigree analysis with clustered markers. Am J Hum Genet 2005; 77(5): 754-767.

Aberg T, Wang X P, Kim J H, Yamashiro T, Bei M, Rice R, Ryoo H M, Thesleff I. Runx2 mediates FGF signaling from epithelium to mesenchyme during tooth morphogenesis. Dev Biol 2004; 270(1): 76-93.

Adzhubei I A, Schmidt S, Peshkin L, Ramensky V E, Gerasimova A, Bork P, Kondrashov A S, Sunyaev S R. A method and server for predicting damaging missense mutations. Nat Methods 2010; 7(4): 248-249.

Altonen M, Myllarniemi S. Results of surgical exposure of impacted cuspids and bicuspids in relation to patients' somatic and dental maturation. Int J Oral Surgery 1976; 5(4): 180-186.

Antonarakis S E, Beckmann J S. Mendelian disorders deserve more attention. Nat Rev Genet 2006; 7(4): 277-282.

Arte S. Phenotypic and Genotypic Features of Familial Hypodontia. 2001. Academic Dissertation, University of Helsinki, Faculty of Medicine, Department of Pedodontics and Orthodontics, Institute of Dentistry.

Arte S, Nieminen P, Pirinen S, Thesleff I, Peltonen L. Gene defect in hypodontia: exclusion of EGF, EGFR, and FGF-3 as candidate genes. J Dent Res 1996; 75(6): 1346-1352.

Baccetti T. [An analysis of the prevalence of isolated dental anomalies and of those associated with hereditary syndromes: a model for evaluating the genetic control of the dentition characteristics]. Minerva Stomatol 1993; 42(6): 281-294.

Baccetti T. A controlled study of associated dental anomalies. Angle Orthod 1998; 68(3): 267-274.

Baccetti T. Tooth anomalies associated with failure of eruption of first and second permanent molars. Am J Orthod Dentofacial Orthop 2000; 118(6): 608-610.

Bamshad M J, Ng S B, Bigham A W, Tabor H K, Emond M J, Nickerson D A, Shendure J. Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet 2011; 12(11): 745-755.

Bass T P. Observations of the misplaced upper canine tooth. Dent Pract (Cincinnati) 1967; 18: 25-33.

Becker A, Smith P, Behar R. The incidence of anomalous maxillary lateral incisors in relation to palatally-displaced cuspids. Angle Orthod 1981; 51(1): 24-29.

Becker A, Zilberman Y, Tsur B. Root length of lateral incisors adjacent to palatally-displaced maxillary cuspids. Angle Orthod 1984; 54(3): 218-225.

Bergendal B, Klar J, Stecksen-Blicks C, Norderyd J, Dahl N. Isolated oligodontia associated with mutations in EDARADD, AXIN2, MSX1, and PAX9 genes. Am J Med Genet A 2011; 155(7): 1616-1622.

Bergland O, Semb G, Abyholm F E. Elimination of the residual alveolar cleft by secondary bone grafting and subsequent orthodontic treatment. Cleft Palate Journal 1986; 23(3): 175-205.

Bezerra M C, Carvalho J F, Prokopowitsch A S, Pereira R M. RANK, RANKL and osteoprotegerin in arthritic bone loss. Braz J Med Biol Res 2005; 38(2): 161-170.

Bjerklin K, Kurol J, Valentin J. Ectopic eruption of maxillary first permanent molars and association with other tooth and developmental disturbances. Eur J Orthod 1992; 14(5): 369-375.

Bloch-Zupan A, Jamet X, Etard C, Laugel V, Muller J, Geoffroy V, Strauss J P, Pelletier V, Marion V, Poch O, Strahle U, Stoetzel C, Dollfus H. Homozygosity mapping and candidate prioritization identify mutations, missed by whole-exome sequencing, in SMOC2, causing major dental developmental defects. Am J Hum Genet 2011; 89(6): 773-781.

Blouet B W. Malta and the Knights of St John. In: The Story Of Malta. Malta: Progress Press, 2004; 6th rev.: 81-122.

Bock C, Reither S, Mikeska T, Paulsen M, Walter J, Lengauer T. BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. Bioinformatics 2005; 21(21): 4067-4068.

Boehnke M. Limits of resolution of genetic linkage studies: implications for the positional cloning of human disease genes. Am J Hum Genet 1994; 55(2): 379-390.

Bot P L, Salmon D. Congenital defects of the upper lateral incisors (ULI): condition and measurements of the other teeth, measurements of the superior arch, head and face. Am J Phys Anthropol 1977; 46(2): 231-243.

Brin I, Becker A, Zilberman Y. Resorbed lateral incisors adjacent to impacted canines have normal crown size. Am J Orthod Dentofacial Orthop 1993; 104(1): 60-66.

Broadbent BH. Ontogenic development of occlusion. Angle Orthod 1941; 11: 223-241.

Bronckers A L, Engelse M A, Cavender A, Gaikwad J, D'Souza R N. Cell-specific patterns of CBFA1 and protein expression in postnatal murine dental tissues. Mech Dev 2001; 101(1-2): 255-258.

Bruder C E, Piotrowski A, Gijsbers A A, Andersson R, Erickson S, Diaz de S T, Menzel U, Sandgren J, von T D, Poplawski A, Crowley M, Crasto C, Partridge E C, Tiwari H, Allison D B, Komorowski J, van Ommen G J, Boomsma D I, Pedersen N L, den Dunnen J T, Wirdefeldt K, Dumanski J P. Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. Am J Hum Genet 2008; 82(3): 763-771.

Cahill D R, Marks S C. Tooth eruption: evidence for the central role of the dental follicle. J Oral Pathol 1980; 9(4): 189-200.

Camilleri S. The prevalence of impacted permanent canines in Maltese schoolchildren - a pilot study. Maltese Medical Journal 1995; 7(1): 42-45.

Camilleri S. Maxillary canine anomalies and tooth agenesis. Eur J Orthod 2005; 27(5): 450-456.

Camilleri S, Lewis CM, McDonald F. Ectopic Maxillary Canines: Segregation Analysis and a Twin Study. J Dent Res 87[6], 580-583. 2008.

Camilleri S, Mulligan K, Pace Balzan A. Dental Association of Malta Scientific Meeting. 1995.

Camilleri S, Mulligan K. The prevalence of malocclusion in Maltese schoolchildren as measured by the Index of Orthodontic Treatment Need. Malta Medical Journal 2007; 19(1): 19-24.

Camm J H, Schuler J L. Premature eruption of the premolars. ASDC J Dent Child 1990; 57(2): 128-133.

Chaushu S, Bongart M, Aksoy A, Ben-Bassat Y, Becker A. Buccal ectopia of maxillary canines with no crowding. Am J Orthod Dentofacial Orthop 2009; 136(2): 218-223.

Chaushu S, Sharabi S, Becker A. Dental morphologic characteristics of normal versus delayed developing dentitions with palatally displaced canines. Am J Orthod Dentofacial Orthop 2002; 121(4): 339-346.

Chaushu S, Sharabi S, Becker A. Tooth size in dentitions with buccal canine ectopia. Eur J Orthod 2003a; 25(5): 485-491.

Chaushu S, Zilberman Y, Becker A. Maxillary incisor impaction and its relationship to canine displacement. Am J Orthod Dentofacial Orthop 2003b; 124(2): 144-150.

Chestnutt I G, Burden D J, Steele J G, Pitts N B, Nuttall N M, Morris A J. The orthodontic condition of children in the United Kingdom, 2003. Br Dent J 2006; 200(11): 609-612.

Choi K Y, Kim H J, Lee M H, Kwon T G, Nah H D, Furuichi T, Komori T, Nam S H, Kim Y J, Kim H J, Ryoo H M. Runx2 regulates FGF2-induced Bmp2 expression during cranial bone development. Dev Dyn 2005; 233(1): 115-121.

Choi M, Scholl U I, Ji W, Liu T, Tikhonova I R, Zumbo P, Nayir A, Bakkaloglu A, Ozen S, Sanjad S, Nelson-Williams C, Farhi A, Mane S, Lifton R P. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A 2009; 106(45): 19096-19101.

Cielinski M J, Iizuka T, Marks S C J. Dental abnormalities in the osteopetrotic rat mutation microphthalmia blanc. Arch Oral Biol 1994; 39(11): 985-990.

Clerget-Darpoux F, Bonaiti-Pellie C, Hochez J. Effects of misspecifying genetic parameters in lod score analysis. Biometrics 1986; 42(2): 393-399.

Clerget-Darpoux F, Elston R C. Are linkage analysis and the collection of family data dead? Prospects for family studies in the age of genome-wide association. Hum Hered 2007; 64(2): 91-96.

Colyer F. The teeth of Londoners in the 17th and 18th Centuries. Dental Record 1922; 42: 237-243.

Coulter J, Richardson A. Normal eruption of the maxillary canine quantified in three dimensions. Eur J Orthod1997; 19(2): 171-183.

Cox K H, DeLeon D V, Angerer L M, Angerer R C. Detection of mrnas in sea urchin embryos by in situ hybridization using asymmetric RNA probes. Dev Biol 1984; 101(2): 485-502.

Craddock N, Hurles M E, Cardin N, Pearson R D, Plagnol V, Robson S, Vukcevic D, Barnes C, Conrad D F, Giannoulatou E, Holmes C, Marchini J L, Stirrups K, Tobin M D, Wain L V, Yau C, Aerts J, Ahmad T, Andrews T D, Arbury H, Attwood A, Auton A, Ball S G, Balmforth A J, Barrett J C, et al. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. Nature 2010; 464(7289): 713-720.

Cui Y, Li G, Li S, Wu R. Designs for linkage analysis and association studies of complex diseases. Methods Mol Biol 2010; 620: 219-242.

Davis P J. Hypodontia and hyperdontia of permanent teeth in Hong Kong schoolchildren. Community Dent Oral Epidemiol 1987; 15(4): 218-220.

Decker E, Stellzig-Eisenhauer A, Fiebig B S, Rau C, Kress W, Saar K, Ruschendorf F, Hubner N, Grimm T, Weber B H. PTHR1 loss-of-function mutations in familial, nonsyndromic primary failure of tooth eruption. Am J Hum Genet 2008; 83(6): 781-786.

Desai J, Shannon M E, Johnson M D, Ruff D W, Hughes L A, Kerley M K, Carpenter D A, Johnson D K, Rinchik E M, Culiat C T. Nell1-deficient mice have reduced expression of extracellular matrix proteins causing cranial and vertebral defects. Hum Mol Genet 2006; 15(8): 1329-1341.

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J. A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 1996; 380(6570): 152-154.

Dipple K M, McCabe E R. Phenotypes of patients with "simple" Mendelian disorders are complex traits: thresholds, modifiers, and systems dynamics. Am J Hum Genet 2000; 66(6): 1729-1735.

Drissi H, Luc Q, Shakoori R, Chuva De Sousa L S, Choi J Y, Terry A, Hu M, Jones S, Neil J C, Lian J B, Stein J L, van Wijnen A J, Stein G S. Transcriptional autoregulation of the bone related CBFA1/RUNX2 gene. J Cell Physiol 2000; 184(3): 341-350.

Ducy P, Starbuck M, Priemel M, Shen J, Pinero G, Geoffroy V, Amling M, Karsenty G. A Cbfa1-dependent genetic pathway controls bone formation beyond embryonic development. Genes Dev 1999; 13(8): 1025-1036.

Dudbridge F. Family-based Assocation. In: Handbook of Statistical Genetics. (Eds. Balding D, Bishop M, Cannings C). John Wiley and Sons, Hoboken, 2007; Third Edition: 1264-1285.

Elliott M, Maher E R. Beckwith-Wiedemann syndrome. J Med Genet 1994; 31(7): 560-564.

Ely N J, Sherriff M, Cobourne M T. Dental transposition as a disorder of genetic origin. Eur J Orthod 2006; 28(2): 145-151.

Enomoto H, Shiojiri S, Hoshi K, Furuichi T, Fukuyama R, Yoshida C A, Kanatani N, Nakamura R, Mizuno A, Zanma A, Yano K, Yasuda H, Higashio K, Takada K, Komori T. Induction of osteoclast differentiation by Runx2 through receptor activator of nuclear factor-kappa B ligand (RANKL) and osteoprotegerin regulation and partial rescue of osteoclastogenesis in Runx2-/- mice by RANKL transgene. J Biol Chem 2003; 278(26): 23971-23977.

Ericson S, Kurol J. Early treatment of palatally erupting maxillary canines by extraction of the primary canines. Eur J Orthod 1988; 10(4): 283-295.

Ericson S, Kurol J. Resorption of incisors after ectopic eruption of maxillary canines: a CT study. Angle Orthod 2000; 70(6): 415-423.

Evangelou E, Trikalinos T A, Salanti G, Ioannidis J P. Family-based versus unrelated case-control designs for genetic associations. PLoS Genet 2006; 2(8): e123.

Evans D M, Cardon L R. Guidelines for genotyping in genomewide linkage studies: single-nucleotide-polymorphism maps versus microsatellite maps. Am J Hum Genet 2004; 75(4): 687-692.

Feichtinger C, Rossiwall B, Wunderer H. Canine transposition as autosomal recessive trait in an inbred kindred. J Dent Res 1977; 56(12): 1449-1452.

Fernandez E, Bravo L A, Canteras M. Eruption of the permanent upper canine: a radiologic study. Am J Orthod Dentofac Orthop1998; 113(4): 414-420.

Feuk L, Carson A R, Scherer S W. Structural variation in the human genome. Nat Rev Genet 2006; 7(2): 85-97.

Franceschi R T, Xiao G. Regulation of the osteoblast-specific transcription factor, Runx2: responsiveness to multiple signal transduction pathways. J Cell Biochem 2003; 88(3): 446-454.

Franceschi R T, Xiao G, Jiang D, Gopalakrishnan R, Yang S, Reith E. Multiple signaling pathways converge on the Cbfa1/Runx2 transcription factor to regulate osteoblast differentiation. Connect Tissue Res 2003; 44 Suppl 1: 109-116.

Frazier-Bowers S A, Guo D C, Cavender A, Xue L, Evans B, King T, Milewicz D, D'Souza R N. A novel mutation in human PAX9 causes molar oligodontia. J Dent Res 2002; 81(2): 129-133.

Gall J G, Pardue M L. Formation and detection of RNA-DNA hybrid molecules in cytological preparations. Proc Natl Acad Sci U S A 1969; 63(2): 378-383.

Gao Y, Kobayashi H, Ganss B. The human KROX-26/ZNF22 gene is expressed at sites of tooth formation and maps to the locus for permanent tooth agenesis (He-Zhao deficiency). J Dent Res 2003; 82(12): 1002-1007.

Gao Y H, Shinki T, Yuasa T, Kataoka-Enomoto H, Komori T, Suda T, Yamaguchi A. Potential role of cbfa1, an essential transcriptional factor for osteoblast differentiation, in osteoclastogenesis: regulation of mRNA expression of osteoclast differentiation factor (ODF). Biochemical and biophysical research communications 1998; 252(3): 697-702.

Garn S, Lewis A, Kerewsky R. Third molar agenesis and variation in size of the remaining teeth. Nature 1964; 201: 839.

Garn S, Lewis A, Vicinus J. Third molar polymorhism and it's significance to dental genetics. J Dent Res 1963; 42: 1344-1363.

Garn S M, Lewis A B. The gradient and the pattern of crown-size reduction in simple hypodontia. Angle Orthod 1970; 40(1): 51-58.

Garyali P, Siwach P, Singh P K, Puri R, Mittal S, Sengupta S, Parihar R, Ganesh S. The malin-laforin complex suppresses the cellular toxicity of misfolded proteins by promoting their degradation through the ubiquitin-proteasome system. Hum Mol Genet 2009; 18(4): 688-700.

Geller F, Feenstra B, Zhang H, Shaffer J R, Hansen T, Esserlind A L, Boyd H A, Nohr E A, Timpson N J, Fatemifar G, Paternoster L, Evans D M, Weyant R J, Levy S M, Lathrop M, Smith G D, Murray J C, Olesen J, Werge T, Marazita M L, Sorensen T I, Melbye M. Genome-wide association study identifies four Loci associated with eruption of permanent teeth. PLoS Genet 2011; 7(9): e1002275.

Geoffroy V, Kneissel M, Fournier B, Boyde A, Matthias P. High Bone Resorption in Adult Aging Transgenic Mice Overexpressing Cbfa1/Runx2 in Cells of the Osteoblastic Lineage. Mol Cell Biol 2002; 22(17): 6222-6233.

Gilissen C, Hoischen A, Brunner H G, Veltman J A. Disease gene identification strategies for exome sequencing. Eur J Hum Genet 2012; 20(5): 490-497.

Giuliana G, Trisi P, D'Arpa M, Scarano A. Cementum growth in impacted teeth. Acta Stomatologica Belgica 1995; 92(1): 7-11.

Goldenberg M, Das P, Messersmith M, Stockton D W, Patel P I, D'Souza R N. Clinical, radiographic, and genetic evaluation of a novel form of autosomal-dominant oligodontia. J Dent Res 2000; 79(7): 1469-1475.

Goldin L R, Weeks D E. Two-locus models of disease: comparison of likelihood and nonparametric linkage methods. Am J Hum Genet 1993; 53(4): 908-915.

Gori F, Hofbauer L C, Dunstan C R, Spelsberg T C, Khosla S, Riggs B L. The expression of osteoprotegerin and RANK ligand and the support of osteoclast formation by stromal-osteoblast lineage cells is developmentally regulated. Endocrinology 2000; 141(12): 4768-4776.

Greenberg D A, Abreu P, Hodge S E. The power to detect linkage in complex disease by means of simple LOD-score analyses. Am J Hum Genet 1998; 63(3): 870-879.

Gropman A L, Adams D R. Atypical patterns of inheritance. Semin Pediatr Neurol 2007; 14(1): 34-45.

Halsey T A, Yang L, Walker J R, Hogenesch J B, Thomas R S. A functional map of NFkappaB signaling identifies novel modulators and multiple system controls. Genome Biol 2007; 8(6): R104.

Hasstedt S J. jPAP: Document-driven software for genetic analysis. Genet Epidemiol 2005; 29: 255.

Heinrich J, Bsoul S, Barnes J, Woodruff K, Abboud S. CSF-1, RANKL and OPG regulate osteoclastogenesis during murine tooth eruption. Arch Oral Biol 2005; 50(10): 897-908.

Huang H, Wise G E. Delay of tooth eruption in null mice devoid of the type I IL-1R gene. European Journal of Oral Sciences 2000; 108(4): 297-302.

Ioannidis J P, Ntzani E E, Trikalinos T A, Contopoulos-Ioannidis D G. Replication validity of genetic association studies. Nat Genet 2001; 29(3): 306-309.

Iseri H, Uzel I. Impaction of maxillary canines and congenitally missing third molars. Description of an ancient skull (7250-6700 BC). Eur J Orthod 1993; 15(1): 1-5.

Jabs E W, Muller U, Li X, Ma L, Luo W, Haworth I S, Klisak I, Sparkes R, Warman M L, Mulliken J B. A mutation in the homeodomain of the human MSX2 gene in a family affected with autosomal dominant craniosynostosis. Cell 1993; 75(3): 443-450.

Jacoby H. The etiology of maxillary canine impactions. Am J Orthod 1983; 84(2): 125-132.

Jiang H, Sodek J, Karsenty G, Thomas H, Ranly D, Chen J. Expression of core binding factor Osf2/Cbfa-1 and bone sialoprotein in tooth development. Mech Dev 1999; 81(1-2): 169-173.

John S, Shephard N, Liu G, Zeggini E, Cao M, Chen W, Vasavda N, Mills T, Barton A, Hinks A, Eyre S, Jones K W, Ollier W, Silman A, Gibson N, Worthington J, Kennedy G C. Whole-genome scan, in a complex disease, using 11,245 single-nucleotide polymorphisms: comparison with microsatellites. Am J Hum Genet 2004; 75(1): 54-64.

Jorde L B. Linkage disequilibrium as a gene-mapping tool. Am J Hum Genet 1995; 56(1): 11-14.

Jumlongras D, Bei M, Stimson J M, Wang W F, DePalma S R, Seidman C E, Felbor U, Maas R, Seidman J G, Olsen B R. A nonsense mutation in MSX1 causes Witkop syndrome. Am J Hum Genet 2001; 69(1): 67-74.

Jumlongras D, Lin J Y, Chapra A, Seidman C, Seidman J, Maas R, Olsen B. A novel missense mutation in the paired domain of PAX9 causes non-syndromic oligodontia. Hum Genet 2004; 114(3): 242-249.

Kaloust S, Ishii K, Vargervik K. Dental development in Apert syndrome. Cleft Palate Craniofac J 1997; 34(2): 117-121.

Kearns M, Preis J, McDonald M, Morris C, Whitelaw E. Complex patterns of inheritance of an imprinted murine transgene suggest incomplete germline erasure. Nucleic Acids Res 2000; 28(17): 3301-3309.

Khoury M, Beaty T, Cohen B. Genetic approaches to familial agregation: II. Segregation analysis. In: Fundamentals of genetic epidemiology. Oxford University Press, New York, 1993; v. 19: 233-283.

Kim B G, Kim H J, Park H J, Kim Y J, Yoon W J, Lee S J, Ryoo H M, Cho J Y. Runx2 phosphorylation induced by fibroblast growth factor-2/protein kinase C pathways. Proteomics 2006; 6(4): 1166-1174.

Kishimoto K, Kitazawa R, Kurosaka M, Maeda S, Kitazawa S. Expression profile of genes related to osteoclastogenesis in mouse growth plate and articular cartilage. Histochem Cell Biol 2006; 125(5): 593-602.

Kodama H, Yamasaki A, Nose M, Niida S, Ohgame Y, Abe M, Kumegawa M, Suda T. Congenital osteoclast deficiency in osteopetrotic (op/op) mice is cured by injections of macrophage colony-stimulating factor. Jour Exp Med 1991; 173(1): 269-272.

Komori T, Yagi H, Nomura S, Yamaguchi A, Sasaki K, Deguchi K, Shimizu Y, Bronson R T, Gao Y H, Inada M, Sato M, Okamoto R, Kitamura Y, Yoshiki S, Kishimoto T. Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. Cell 1997; 89(5): 755-764.

Kotsomitis N, Freer T J. Inherited dental anomalies and abnormalities. ASDC J Dent Child 1997; 64(6): 405-408.

Kruglyak L. Genetic isolates: separate but equal? Proc Natl Acad Sci U S A 1999; 96(4): 1170-1172.

Kruglyak L, Daly M J, Reeve-Daly M P, Lander E S. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 1996; 58(6): 1347-1363.

Ku C S, Naidoo N, Pawitan Y. Revisiting Mendelian disorders through exome sequencing. Hum Genet 2011; 129(4): 351-370.

Kumar P, Henikoff S, Ng P C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009; 4(7): 1073-1081.

Lammi L, Arte S, Somer M, Jarvinen H, Lahermo P, Thesleff I, Pirinen S, Nieminen P. Mutations in AXIN2 cause familial tooth agenesis and predispose to colorectal cancer. Am J Hum Genet 2004; 74(5): 1043-1050.

Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 1995; 11(3): 241-247.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009; 25(16): 2078-2079.

Lidral A C, Reising B C. The role of MSX1 in human tooth agenesis. J Dent Res 2002; 81(4): 274-278.

Liu W, Toyosawa S, Furuichi T, Kanatani N, Yoshida C, Liu Y, Himeno M, Narai S, Yamaguchi A, Komori T. Overexpression of Cbfa1 in osteoblasts inhibits osteoblast maturation and causes osteopenia with multiple fractures. J Cell Biol 2001; 155(1): 157-166.

Loevy H T. The effect of primary tooth extraction on the eruption of succedaneous premolars. J Am Dent Assoc 1989; 118(6): 715-718.

MacLean C J, Bishop D T, Sherman S L, Diehl S R. Distribution of lod scores under uncertain mode of inheritance. Am J Hum Genet 1993; 52(2): 354-361.

Mamanova L, Coffey A J, Scott C E, Kozarewa I, Turner E H, Kumar A, Howard E, Shendure J, Turner D J. Target-enrichment strategies for next-generation sequencing. Nat Methods 2010; 7(2): 111-118.

Mangion J, Rahman N, Edkins S, Barfoot R, Nguyen T, Sigurdsson A, Townend J V, Fitzpatrick D R, Flanagan A M, Stratton M R. The gene for cherubism maps to chromosome 4p16.3. Am J Hum Genet 1999; 65(1): 151-157.

Marks S C, Cahill D R, Wise G E. The cytology of the dental follicle and adjacent alveolar bone during tooth eruption in the dog. Am J Anat 1983; 168(3): 277-289.

Marks S C J, Cahill D R. Experimental study in the dog of the non-active role of the tooth in the eruptive process. Arch Oral Biol1984; 29(4): 311-322.

Marks S C J, Cahill D R. Regional control by the dental follicle of alterations in alveolar bone metabolism during tooth eruption. J Oral Pathol 1987; 16(4): 164-169.

Marks S C J, Schroeder H E. Tooth eruption: theories and facts. Anat Rec 1996; 245(2): 374-393.

Mathe E, Olivier M, Kato S, Ishioka C, Hainaut P, Tavtigian S V. Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods. Nucleic Acids Res 2006; 34(5): 1317-1325.

McCarroll S A. Extending genome-wide association studies to copy-number variation. Hum Mol Genet 2008; 17(R2): R135-R142.

McCarroll S A, Kuruvilla F G, Korn J M, Cawley S, Nemesh J, Wysoker A, Shapero M H, de Bakker P I, Maller J B, Kirby A, Elliott A L, Parkin M, Hubbell E, Webster T, Mei R, Veitch J, Collins P J, Handsaker R, Lincoln S, Nizzari M, Blume J, Jones K W, Rava R, Daly M J, Gabriel S B, Altshuler D. Integrated detection and population-genetic analysis of SNPs and copy number variation. Nat Genet 2008; 40(10): 1166-1174.

McKay C. The unerupted maxillary canine. An assessment of the role of surgery in 2,500 treated cases. Br Dent J 1978; 145(7): 207-210.

McSherry P, Richardson A. Ectopic eruption of the maxillary canine quantified in three dimensions on cephalometric radiographs between the ages of 5 and 15 years. Eur J Orthod1999; 21(1): 41-48.

Mei R, Galipeau P C, Prass C, Berno A, Ghandour G, Patil N, Wolff R K, Chee M S, Reid B J, Lockhart D J. Genome-wide detection of allelic imbalance using human SNPs and high-density DNA arrays. Genome Res 2000; 10(8): 1126-1137.

Ming J E, Muenke M. Multiple hits during early embryonic development: digenic diseases and holoprosencephaly. Am J Hum Genet 2002; 71(5): 1017-1032.

Mortazavi A, Leeper Thompson E C, Garcia S T, Myers R M, Wold B. Comparative genomics modeling of the NRSF/REST repressor network: from single conserved sites to genome-wide repertoire. Genome Res 2006; 16(10): 1208-1221.

Morton N. Sequential tests for the detection of linkage. Am J Hum Genet 1955; 7(3): 277-318.

Moss J P. The unerupted canine. Dent Pract Dent Rec 1972; 22(6): 241-248.

Mostowska A, Kobielak A, Biedziak B, Trzeciak W H. Novel mutation in the paired box sequence of PAX9 gene in a sporadic form of oligodontia. Eur J Oral Sci 2003; 111(3): 272-276.

Mourshed F. A roentgenographic study of dentigerous cysts. II. Role of roentgenograms in detecting dentigerous cyst in the early stages. Oral Surg Oral Med Oral Pathol 1964; 18: 54-61.

Mundlos S, Otto F, Mundlos C, Mulliken J B, Aylsworth A S, Albright S, Lindhout D, Cole W G, Henn W, Knoll J H, Owen M J, Mertelsmann R, Zabel B U, Olsen B R. Mutations involving the transcription factor CBFA1 cause cleidocranial dysplasia. Cell 1997; 89(5): 773-779.

Nelson G C. Maxillary canine/third premolar transposition in a prehistoric population from Santa Cruz Island, California. Am J Phys Anthropol 1992; 88(2): 135-144.

Neubuser A, Peters H, Balling R, Martin G R. Antagonistic interactions between FGF and BMP signaling pathways: a mechanism for positioning the sites of tooth formation. Cell 1997; 90(2): 247-255.

Ng S B, Bigham A W, Buckingham K J, Hannibal M C, McMillin M J, Gildersleeve H I, Beck A E, Tabor H K, Cooper G M, Mefford H C, Lee C, Turner E H, Smith J D, Rieder M J, Yoshiura K, Matsumoto N, Ohta T, Niikawa N, Nickerson D A, Bamshad M J, Shendure J. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. Nat Genet 2010a; 42(9): 790-793.

Ng S B, Buckingham K J, Lee C, Bigham A W, Tabor H K, Dent K M, Huff C D, Shannon P T, Jabs E W, Nickerson D A, Shendure J, Bamshad M J. Exome sequencing identifies the cause of a mendelian disorder. Nat Genet 2010b; 42(1): 30-35.

Ng S B, Turner E H, Robertson P D, Flygare S D, Bigham A W, Lee C, Shaffer T, Wong M, Bhattacharjee A, Eichler E E, Bamshad M, Nickerson D A, Shendure J. Targeted capture and massively parallel sequencing of 12 human exomes. Nature 2009; 461(7261): 272-276.

Nieminen P, Arte S, Pirinen S, Peltonen L, Thesleff I. Gene defect in hypodontia: exclusion of MSX1 and MSX2 as candidate genes. Hum Genet 1995; 96(3): 305-308.

Nieminen P, Arte S, Tanner D, Paulin L, Alaluusua S, Thesleff I, Pirinen S. Identification of a nonsense mutation in the PAX9 gene in molar oligodontia. Eur J Hum Genet 2001; 9(10): 743-746.

Niu D F, Kondo T, Nakazawa T, Oishi N, Kawasaki T, Mochizuki K, Yamane T, Katoh R. Transcription factor Runx2 is a regulator of epithelial-mesenchymal transition and invasion in thyroid carcinomas. Lab Invest 2012; 92(8): 1181-1190.

Nordenram A, Stromberg C. Positional variations of the impacted upper canine. A clinical and radiologic study. Oral Surg Oral Med Oral Pathol 1966; 22(6): 711-714.

Notini A J, Craig J M, White S J. Copy number variation and mosaicism. Cytogenet Genome Res 2008; 123(1-4): 270-277.

Notoya M, Otsuka E, Yamaguchi A, Hagiwara H. Runx-2 is not essential for the vitamin D-regulated expression of RANKL and osteoprotegerin in osteoblastic cells. Biochem Biophys Res Commun 2004; 324(2): 655-660.

O'Meara W F. Effect of primary molar extraction on gingival emergence of succedaneous tooth. J Dent Res1966; 45(4): 1174-1183.

Ohazama A, Courtney J M, Sharpe P T. Opg, Rank, and Rankl in tooth development: co-ordination of odontogenesis and osteogenesis. J Dent Res 2004; 83(3): 241-244.

Oliver R G, Mannion J E, Robinson J M. Morphology of the maxillary lateral incisor in cases of unilateral impaction of the maxillary canine. Br J Orthod1989; 16(1): 9-16.

Ott J, Kamatani Y, Lathrop M. Family-based designs for genome-wide association studies. Nat Rev Genet 2011; 12(7): 465-474.

Parkin N, Benson P E, Shah A, Thind B, Marshman Z, Glenroy G, Dyer F. Extraction of primary (baby) teeth for unerupted palatally displaced permanent canine teeth in children. Cochrane Database Syst Rev 2009;(2): CD004621.

Paschos E, Huth K C, Fassler H, Rudzki-Janson I. Investigation of maxillary tooth sizes in patients with palatal canine displacement. J Orofac Orthop 2005; 66(4): 288-298.

Peck L, Peck S, Attia Y. Maxillary canine-first premolar transposition, associated dental anomalies and genetic basis. Angle Orthod 1993; 63(2): 99-109.

Peck S, Peck L, Kataja M. The palatally displaced canine as a dental anomaly of genetic origin. Angle Orthod 1994; 64(4): 249-256.

Peck S, Peck L, Kataja M. Concomitant occurrence of canine malposition and tooth agenesis: evidence of orofacial genetic fields. Am J Orthod Dentofacial Orthop 2002; 122(6): 657-660.

Pereira R, Halford K, Sokolov B P, Khillan J S, Prockop D J. Phenotypic variability and incomplete penetrance of spontaneous fractures in an inbred strain of transgenic mice expressing a mutated collagen gene (COL1A1). J Clin Invest 1994; 93(4): 1765-1769.

Pericak-Vance M A, Haines J L. Approaches to gene mapping in complex human diseases. New York ; Chichester : Wiley-Liss, 1998.

Pillas D, Hoggart C J, Evans D M, O'Reilly P F, Sipila K, Lahdesmaki R, Millwood I Y, Kaakinen M, Netuveli G, Blane D, Charoen P, Sovio U, Pouta A, Freimer N, Hartikainen A L, Laitinen J, Vaara S, Glaser B, Crawford P, Timpson N J, Ring S M, Deng G, Zhang W, McCarthy M I, Deloukas P, et. al. Genome-wide association study reveals multiple loci associated with primary tooth development during infancy. PLoS Genet 2010; 6(2): e1000856.

Pirinen S, Arte S, Apajalahti S. Palatal displacement of canine is genetic and related to congenital absence of teeth. J Dent Res 1996; 75(10): 1742-1746.

Power S M, Short M B. An investigation into the response of palatally displaced canines to the removal of deciduous canines and an assessment of factors contributing to favourable eruption. Br J Orthod 1993; 20(3): 215-223.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M A, Bender D, Maller J, Sklar P, de Bakker P I, Daly M J, Sham P C. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81(3): 559-575.

Qiu M, Bulfone A, Martinez S, Meneses J J, Shimamura K, Pedersen R A, Rubenstein J L. Null mutation of Dlx-2 results in abnormal morphogenesis of proximal first and second branchial arch derivatives and abnormal differentiation in the forebrain. Genes and Development 1995; 9(20): 2523-2538.

Quinlan A R, Hall I M. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010; 26(6): 841-842.

Rakyan V K, Preis J, Morgan H D, Whitelaw E. The marks, mechanisms and memory of epigenetic states in mammals. Biochem J 2001; 356(Pt 1): 1-10.

Reich D E, Cargill M, Bolk S, Ireland J, Sabeti P C, Richter D J, Lavery T, Kouyoumjian R, Farhadian S F, Ward R, Lander E S. Linkage disequilibrium in the human genome. Nature 2001; 411(6834): 199-204.

Richard I, Broux O, Allamand V, Fougerousse F, Chiannilkulchai N, Bourg N, Brenguier L, Devaud C, Pasturaud P, Roudaut C, . Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. Cell 1995; 81(1): 27-40.

Risch N. Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 1990; 46(2): 222-228.

Risch N. A note on multiple testing procedures in linkage analysis. Am J Hum Genet 1991; 48(6): 1058-1064.

Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996; 273(5281): 1516-1517.

Robey P G, Fedarko N S, Hefferan T E, Bianco P, Vetter U K, Grzesik W, Friedenstein A, Van der P G, Mintz K P, Young M F, . Structure and molecular regulation of bone matrix proteins. J Bone Miner Res 1993; 8 Suppl 2: S483-S487.

Robinson P N, Krawitz P, Mundlos S. Strategies for exome and genome sequence data analysis in disease-gene discovery projects. Clin Genet 2011; 80(2): 127-132.

Rose J S. A survey of congenitally missing teeth, excluding third molars, in 6000 orthodontic patients. Dent Pract (Cincinnati) 1966; 17(3): 107-114.

Rossini L, Hashimoto Y, Suzuki H, Kurita M, Gianfriddo M, Scali C, Roncarati R, Franceschini D, Pollio G, Trabalzini L, Terstappen G C, Matsuoka M, Caricasole A. VSTM2L is a novel secreted antagonist of the neuroprotective peptide Humanin. FASEB J 2011; 25(6): 1983-2000.

Ruschendorf F, Nurnberg P. ALOHOMORA: a tool for linkage analysis using 10K SNP array data. Bioinformatics 2005; 21(9): 2123-2125.

Sacerdoti R, Baccetti T. Dentoskeletal features associated with unilateral or bilateral palatal displacement of maxillary canines. Angle Orthod 2004; 74(6): 725-732.

Sandovici I, Leppert M, Hawk P R, Suarez A, Linares Y, Sapienza C. Familial aggregation of abnormal methylation of parental alleles at the IGF2/H19 and IGF2R differentially methylated regions. Hum Mol Genet 2003; 12(13): 1569-1578.

Savona-Ventura C, Grech E S. Multiple pregnancy in the Maltese population. Int J Gynaecol Obstet 1988; 26(1): 41-50.

Schroeder H E, Luder H U, Bosshardt D D. Morphological and labeling evidence supporting and extending a modern theory of tooth eruption. Schweizer Monatsschrift fur Zahnmedizin 1992; 102(1): 20-31.

Sellick G S, Longman C, Tolmie J, Newbury-Ecob R, Geenhalgh L, Hughes S, Whiteford M, Garrett C, Houlston R S. Genomewide linkage searches for Mendelian disease loci can be efficiently conducted using high-density SNP genotyping arrays. Nucleic Acids Res 2004; 32(20): e164.

Shalish M, Chaushu S, Wasserstein A. Malposition of unerupted mandibular second premolar in children with palatally displaced canines. Angle Orthod 2009; 79(4): 796-799.

Shalish M, Peck S, Wasserstein A, Peck L. Increased occurrence of dental anomalies associated with infraocclusion of deciduous molars. Angle Orthod 2010; 80(3): 440-445.

Shevde N K, Bendixen A C, Maruyama M, Li B L, Billmire D A. Enhanced activity of osteoblast differentiation factor (PEBP2alphaA2/CBFa1) in affected sutural osteoblasts from patients with nonsyndromic craniosynostosis. Cleft Palate Craniofac J 2001; 38(6): 606-614.

Shiio Y, Rose D W, Aur R, Donohoe S, Aebersold R, Eisenman R N. Identification and characterization of SAP25, a novel component of the mSin3 corepressor complex. Mol Cell Biol 2006; 26(4): 1386-1397.

Silva S, Martins Y, Matias A, Blickstein I. Why are monozygotic twins different? J Perinat Med 2011; 39(2): 195-202.

Singh S M, Murphy B, O'Reilly R. Epigenetic contributors to the discordance of monozygotic twins. Clin Genet 2002; 62(2): 97-103.

Singleton A B. Exome sequencing: a transformative technology. Lancet Neurol 2011; 10(10): 942-946.

Smith K R, Bromhead C J, Hildebrand M S, Shearer A E, Lockhart P J, Najmabadi H, Leventer R J, McGillivray G, Amor D J, Smith R J, Bahlo M. Reducing the exome

search space for mendelian diseases using genetic linkage analysis of exome genotypes. Genome Biol 2011; 12(9): R85.

Stockton D W, Das P, Goldenberg M, D'Souza R N, Patel P I. Mutation of PAX9 is associated with oligodontia. Nat Genet 2000; 24(1): 18-19.

Suzuki T, Suda N, Ohyama K. Osteoclastogenesis during mouse tooth germ development is mediated by receptor activator of NFKappa-B ligand (RANKL). J Bone Miner Metab 2004; 22(3): 185-191.

Svinhufvud E, Myllarniemi S, Norio R. Dominant inheritance of tooth malpositions and their association to hypodontia. Clin Genet 1988; 34(6): 373-381.

Thiele H, Nurnberg P. HaploPainter: a tool for drawing pedigrees with complex haplotypes. Bioinformatics 2005; 21(8): 1730-1732.

Thilander B, Jakobsson S O. Local factors in impaction of maxillary canines. Acta Odont Scand 1968; 26(2): 145-168.

Thirunavukkarasu K, Halladay D L, Miles R R, Yang X, Galvin R J, Chandrasekhar S, Martin T J, Onyia J E. The osteoblast-specific transcription factor Cbfa1 contributes to the expression of osteoprotegerin, a potent inhibitor of osteoclast differentiation and function. J Bioll Chem 2000; 275(33): 25163-25172.

Tiziani V, Reichenberger E, Buzzo C L, Niazi S, Fukai N, Stiller M, Peters H, Salzano F M, Raposo do Amaral C M, Olsen B R. The gene for cherubism maps to chromosome 4p16. Am J Hum Genet 1999; 65(1): 158-166.

Tsutsumi S, Kamata N, Vokes T J, Maruoka Y, Nakakuki K, Enomoto S, Omura K, Amagasa T, Nagayama M, Saito-Ohara F, Inazawa J, Moritani M, Yamaoka T, Inoue H, Itakura M. The novel gene encoding a putative transmembrane protein is mutated in gnathodiaphyseal dysplasia (GDD). Am J Hum Genet 2004; 74(6): 1255-1261.

Ueki Y, Tiziani V, Santanna C, Fukai N, Maulik C, Garfinkle J, Ninomiya C, doAmaral C, Peters H, Habal M, Rhee-Morris L, Doss J B, Kreiborg S, Olsen B R, Reichenberger E. Mutations in the gene encoding c-Abl-binding protein SH3BP2 cause cherubism. Nat Genet 2001; 28(2): 125-126.

van den Boogaard M J, Dorland M, Beemer F A, van Amstel H K. MSX1 mutation is associated with orofacial clefting and tooth agenesis in humans. Nat Genet 2000; 24(4): 342-343.

Van Heyningen V, Yeyati P L. Mechanisms of non-Mendelian inheritance in genetic disease. Hum Mol Genet 2004; 13 Spec No 2: R225-R233.

Vastardis H, Karimbux N, Guthua S W, Seidman J G, Seidman C E. A human MSX1 homeodomain missense mutation causes selective tooth agenesis. Nat Genet 1996; 13(4): 417-421.

Vissers L E, de L J, Gilissen C, Janssen I, Steehouwer M, de V P, van L B, Arts P, Wieskamp N, del R M, van Bon B W, Hoischen A, de Vries B B, Brunner H G, Veltman J A. A de novo paradigm for mental retardation. Nat Genet 2010; 42(12): 1109-1112.

Voronkov A, Krauss S. Wnt/beta-catenin signaling and small molecule inhibitors. Curr Pharm Des 2013; 19(4): 634-664.

Wang J, Jian F, Chen J, Wang H, Lin Y, Yang Z, et al. Sequence analysis of PAX9, MSX1 and AXIN2 genes in a Chinese oligodontia family. Arch Oral Biol 2011;56(10):1027-34.

Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010; 38(16): e164.

Wang Q, Peng P, Qian M, Wan L, Deng M. Hybridization and amplification rate correction for affymetrix SNP arrays. BMC Med Genomics 2012; 5: 24.

Wei Q, Wang L, Wang Q, Kruger W D, Dunbrack R L, Jr. Testing computational prediction of missense mutation phenotypes: functional characterization of 204 mutations of human cystathionine beta synthase. Proteins 2010; 78(9): 2058-2074.

Weinberg C R. Methods for detection of parent-of-origin effects in genetic studies of case-parents triads. Am J Hum Genet 1999; 65(1): 229-235.

Wenzel K, Daskalow K, Herse F, Seitz S, Zacharias U, Schenk J A, Schulz H, Hubner N, Micheel B, Schlag P M, Osterziel K J, Ozcelik C, Scherneck S, Jandrig B. Expression of the protein phosphatase 1 inhibitor KEPI is downregulated in breast cancer cell lines and tissues and involved in the regulation of the tumor suppressor EGR1 via the MEK-ERK pathway. Biol Chem 2007; 388(5): 489-495.

Westendorf J J. Transcriptional co-repressors of Runx2. J Cell Biochem 2006; 98(1): 54-64.

Wilcox J N. Fundamental principles of in situ hybridization. J Histochem Cytochem 1993; 41(12): 1725-1733.

Wilkie A O M, Morriss-Kay G M. Genetics of Craniofacial Development and Malformation. Nat Rev Genet 2001; 2(6): 458-468.

Wilson A F, Ziegler A. Lessons learned from Genetic Analysis Workshop 17: transitioning from genome-wide association studies to whole-genome statistical genetic analysis. Genet Epidemiol 2011; 35 Suppl 1: S107-S114.

Wise G E, Ding D, Yao S. Regulation of secretion of osteoprotegerin in rat dental follicle cells. Eur J Oral Sci 2004; 112(5): 439-444.

Wise G E, Frazier-Bowers S, D'Souza R N. Cellular, molecular, and genetic determinants of tooth eruption. Crit Rev Oral Biol Med 2002; 13(4): 323-334.

Wise G E, Lumpkin S J, Huang H, Zhang Q. Osteoprotegerin and osteoclast differentiation factor in tooth eruption. J Dent Res 2000; 79(12): 1937-1942.

Wise G E, Yao S. Regional differences of expression of bone morphogenetic protein-2 and RANKL in the rat dental follicle. Eur J Oral Sci 2006; 114(6): 512-516.

Wise G E, Yao S, Henk W G. Bone formation as a potential motive force of tooth eruption in the rat molar. Clinical anatomy (New York, N Y) 2007; 20(6): 632-639.

Wise G E, Yao S, Odgren P R, Pan F. CSF-1 regulation of osteoclastogenesis for tooth eruption. J Dent Res 2005; 84(9): 837-841.

Yamaguchi T, Hosomichi K, Narita A, Shirota T, Tomoyasu Y, Maki K, Inoue I. Exome resequencing combined with linkage analysis identifies novel PTH1R variants in primary failure of tooth eruption in Japanese. J Bone Miner Res 2011; 26(7): 1655-1661.

Yilmaz H H, Turkkahraman H, Sayin M O. Prevalence of tooth transpositions and associated dental anomalies in a Turkish population. Dentomaxillofac Radiol 2005; 34(1): 32-35.

Yoda S, Suda N, Kitahara Y, Komori T, Ohyama K. Delayed tooth eruption and suppressed osteoclast number in the eruption pathway of heterozygous Runx2/Cbfa1 knockout mice. Arch Oral Biol 2004; 49(6): 435-442.

Young M F, Kerr J M, Ibaraki K, Heegaard A M, Robey P G. Structure, expression, and regulation of the major noncollagenous matrix proteins of bone. Clin Orthop Relat Res 1992;(281): 275-294.

Zilberman Y, Cohen B, Becker A. Familial trends in palatal canines, anomalous lateral incisors, and related phenomena. Eur J Orthod 1990; 12(2): 135-139.

Zou S J, D'Souza R N, Ahlberg T, Bronckers A L. Tooth eruption and cementum formation in the Runx2/Cbfa1 heterozygous mouse. Arch Oral Biol 2003; 48(9): 673-677.

Appendix 1. Ethical Approval

L-UNIVERSITÀ TA' MALTA
Msida - Malta
SKOLA MEDIKA

UNIVERSITY OF MALTA
Msida - Malta
MEDICAL SCHOOL

REF. TAGHNA: *OUR REF.:*

REF. TIEGHEK: *YOUR REF.:*

Ref No: 063 /2004

29th July 2004

Dr S Camilleri
Dept of Dental Surgery
Medical School
G'Mangia

Dear Dr Camilleri

Please refer to your application submitted to the Research Ethics Committee in connection with your clinical research entitled:

THE AETIOLOGY OF ECTOPIC MAXILLARY CANINE TEETH

At the last meeting of the Research Ethics Committee held on 19th July 2004 members reviewed and approved the above-mentioned Protocol.

You are kindly requested to submit to the Research Ethics Committee a brief report on completion of your research.

Yours sincerely

Dr. R. Golea.

Professor A Fenech
Chairman
Research Ethics Committee

Cc Supervisor

TEL: (356) 21333903-6, 21336451 CABLES: UNIVERSITY-MALTA TELEX: 407 HIEDUC MW FAX: (356) 21336450
MEDICAL SCHOOL, GWARDAMANGIA TEL: 21221019, 21239783 FAX: (010 356) 21235638

L-UNIVERSITÀ TA' MALTA
Msida - Malta
SKOLA MEDIKA

UNIVERSITY OF MALTA
Msida - Malta
MEDICAL SCHOOL

REF. TAGHNA:

REF. TIEGHEK:

OUR REF.:

YOUR REF.:

Ref No: 05/2007

3rd April 2007

Dr Simon Camilleri
Faculty of Dental Surgery
Medical School
G'Mangia

Dear Dr Camilleri

Please refer to your application submitted to the Research Ethics Committee in connection with your research entitled:

THE AETIOLOGY OF ECTOPIC MAXILLARY CANINE TEETH

The University Research Ethics Committee at its meeting of 30th March 2007 approved the above-mentioned Protocol.

Yours sincerely

Dr M Vassallo
Chairman
Research Ethics Committee

TEL: (356) 21333903-6, 21336451 CABLES: UNIVERSITY-MALTA TELEX: 407 HIEDUC MW FAX: (356) 21336450
MEDICAL SCHOOL, GWARDAMANGIA TEL: 21221019, 21239783 FAX: (010 356) 21235638

CONSENT FORM

I am a Maltese citizen and am over/under* eighteen (18) years of age.

I have been asked to participate in a research study entitled: The Aetiology of Ectopic Maxillary Canine Teeth

The purpose and details of the study have been explained to me by Dr S Camilleri and any difficulties which I raised have been adequately clarified.

I give my consent to the Principal Investigator and his delegate either make the appropriate observations/tests or both or take the necessary samples. I am aware of any inconveniences which this will cause.

I understand that the results of this study may be used for medical or scientific purposes and that the results achieved from this study in which I am participating may be reported or published: however, I shall not be personally identified in any way, either individually or collectively, without my express written permission.

I am under no obligation to participate in this study and am doing so voluntarily.

I may withdraw from the study at any time, without giving any reason. This will not influence in any way the care and attention and treatment normally given to me (applicable only in case of patients receiving treatment).

I understand that any complications and/or adverse effects which may arise during or as a consequence of the study will be recorded and any treatment which this may entail will be given within the Government Health Services.

I am not receiving any remuneration for participating in this study.

In case of queries during the study I may contact Dr S Camilleri. on 23401876 or Email: simon.camilleri@um.edu.mt

Signature of participant. (In case of minors, parent or guardian to sign).

Name of participant (in block letters)

Id. No.:

Signature of Chief Investigator/Investigator

Name of Chief Investigator/Investigator

Id. No.: ……..

DATE

.* delete where applicable

FORMOLA TA KUNSENS

Jiena cittadin Malti u qbizt/ma qbiztx* it-tmintax (18) is-sena.

Talbuni biex niehu sehem fi studju ricerka bl-isem ta': Investigazzjoni fi snien ma telghux

Il-ghan u d-dettalji ta' l-istudju spjegahomli it-tabib Simon Camilleri, li wkoll iccarali xi mistoqsijiet li ghamilt.

Naghti l-kunsens tieghi lill-persuna responsabbli ghal din ir-ricerka u l-assistenti taghha biex jaghmlu l-osservazjonijiet li hemm bzonn jew inkella jiehdu l-kampjuni u nifhem li dan jista' jkun ta' skomdu ghalija.

Jiena nifhem li r-rizultati ta' dan l-istudju jistghu jintuzaw ghal skopijiet xjentifici u jista' jigi ppubblikat rapport bil-miktub: jekk isir hekk, b'ebda mod ma nista' nkun identifikat/a, individwalment jew bhala parti minn grupp, minghajr il-kunsens tieghi bil-miktub.

Jiena ma ghandi l-ebda dmir li niehu sehem f'dan l-istudju u dan qed naghmlu minn rajja.

Jiena nista', meta rrid, ma nkomplix niehu sehem fl-istudju, u minghajr ma' naghti raguni. Jekk naghmel hekk xorta nibqa' niehu l-kura li ssoltu tinghatali (tapplika biss ghal pazjenti Ii qed jiehdu kura).

Jiena nifhem li jekk ikun hemm xi kumplikazzjoni jew effetti mhux mistennija waqt l-istudju, dawn jigu mnizzla bil-miktub u jekk ikun hemm bzonn xi kura, kif jinghata fis-Servizz Nazjonali tas-Sahha.

Jiena mhux qed nithallas biex niehu sehem f'dan l-istudju.

Jekk ikolli xi diffikulta' waqt l-istudju, nista' nistaqsi ghal: Simon Camillei fuq numru 23401876 Email: simon.camilleri@um.edu.mt

Firma tal-participant. (Filkas ta’ minorenni, jiffirma il-genitur jew gwardjan.)

Isem tal-participant (b 'ittri kbar)

Numru ta' l-identita

Firma tal-persuna responsabbli ghal din ir-ricerka

Isem tal-persuna responsabbli ghal din ir-ricerka

Numru ta' l-identita’

DATA

*aqta fejn ma japplikax

INFORMATION SHEET

CONSENT TO PARTICIPATE IN A GENETIC STUDY INVESTIGATING UNERUPTED TEETH

DESCRIPTION OF THE STUDY. The University of Malta is carrying out a study on impacted maxillary canine teeth in Maltese nationals and is inviting you to participate in this study. This work would involve compilation of a family tree and isolation, analysis and storage of your DNA (molecule of heredity).

This study is scheduled to continue for a number of years and participation is voluntary. You will be required to supply relevant medical and dental history, together with a sample of cells obtained from the inside of your mouth by means of a mouthwash. This will be used to obtain DNA. The DNA will be used to produce molecular genetic data that will be assembled into a database.

PROCEDURES. You would be asked questions (either directly or by questionnaire) regarding your and your family's general health and dental condition.

You would be provided with one bottle of ordinary mouthwash. You would be required to rinse out your mouth and then spit the mouthwash back into the bottle, once first thing in the morning on rising. The mouthwash would then be collected for examination.

Should it be necessary, we may ask you to have a dental examination, photographs, Xrays or casts of your teeth at our expense.

YOUR RIGHTS. Participation in this study is voluntary. You have the right to end your participation at any time, and to decide whether the material already collected can remain part of the study or must be destroyed. You also have the right to determine which of the above procedures you are willing to complete. Confidentiality will be preserved in any scientific publication of the data, and we will not provide genetic information back to you, except in the unusual circumstances listed in 'OTHER ISSUES', below. You may ask questions at any time during your participation.

RISKS, INCONVENIENCE AND DISCOMFORT. There is no risk to your health, nor are any uncomfortable procedures involved during participation.

BENEFITS. There is no material benefit to you for taking part. However others might benefit in the event of a test or cure being developed as a direct result of this study.

OTHER ISSUES RELATED TO THIS PROPOSAL

FUTURE RESEARCH. Our specific research plans have been summarized in the first paragraph of this document. In the course of the research, we will obtain samples of your cells. While we will do some testing immediately, we may store your DNA from your cells in the freezer for related studies. The storage will be for a maximum period of 6 years, unless the sample is used up before this time. In addition, certain

information obtained will be stored in our research records. Your DNA will not be used for other research purposes without your prior permission. If you agree to the use of your sample in future research, you will be able to discuss any concerns you have about this research with a genetic counselor or your medical doctor.

UNANTICIPATED MEDICAL INFORMATION. During the course of this investigation, it is possible (although not likely) that we will obtain unanticipated information about your health or genetic makeup. If this information is deemed to be vital to your medical health, we will provide this to your physician after having obtained your permission. This information cannot compromise any other individual. Circumstances may require further information or another cell sample. In this case we will contact you through your physician.

FAMILY RELATIONSHIPS. In the course of this study, it is possible (although not likely) that we may discover information about relationships within the family, such as adoption or paternity. We will not provide this type of information to anybody. The exception would be the extraordinary circumstance should this information be required for the medical care of the individuals involved. If we are convinced that this is necessary, we will provide the information to the physician who is providing medical care to the patient.

CONFIDENTIALITY. All information collected or discovered about you or your family is considered confidential and private. All relevant records and files will be kept in a locked office in the University of Malta. Control of this information is in your hands. No-one, not even other members of your family can access this information without your permission.

COLLECTION OF, RESEARCH ON, AND STORAGE OF GENETIC MATERIAL. It should be clear to you that your DNA sample will only be used for research. No other testing or research will be conducted on your sample unless you specifically give permission (as indicated below). The DNA isolated from your cell sample will be stored in locked freezers contained in secure buildings at the University of Malta. The samples will be labeled and stored by codes defined by us. The only individual who will have access to the codes will be the principal investigator of this study, Dr. Simon Camilleri. Dr Camilleri undertakes to exchange only coded anonymised research results with foreign investigators, and therefore your identity will not be revealed to anyone.

CONTACTS. This study is supervised by the Principal Investigator, Dr. Simon Camilleri who can be contacted at 23401876 or email simon.camilleri@um.edu.mt with any related questions.

TAGHRIF

KUNSENS BIEX TIE}U SEHEM FI STUDJU GENETIKU - INVESTIGAZZJONI TA' SNIEN LI MA TELGHUX.

TIFSIRA TAL-ISTUDJU. L-Universita' ta' Malta qed taghmel studju fuq certi snien (imsejhin in-nejbiet ta'fuq) li ma jitilghux. Dan l-istudju qed isir fuq nies ta'nazzjonalita Maltija u inti qed tigi mistieden/na tiehu sehem. Dan jinvolvi kompilazzjoni tal- '*family tree'* tieghek u isolazzjoni, analisi u hazna tad-DNA, li hija il-molekula tal-eredita' tieghek.

Dan l-istudju mahsub li jkompli ghal xi 6 snin. Is-sehem huwa volontarju. Inkunu nixtiequ min ghandek xi informazzjoni dwar sahhtek generali u ta' snienek flimkien ma ftit celloli min halqek. Ic-celloli jigu migbura bit-tlahlih ta' halqek. Min dawn ic-celloli niehdu id-DNA. Minnu nohorgu *data* fuq l-istruttura genetika li jigi imdahhal fc*database*

PROCEDURA. Tigi mitlub/a informazzjoni dwar is-sahha generali w is-sahha ta' snienek u tal-familja tieghek.

Tigi mitlub kampjun bzieq billi tobzoq gol-flixkun ipprovdut. Dan jigi migbur u ezaminat.

Jekk ikun hemm bzonn, nitolbu li jsir ezami ta snienek jew niehdu ritratti, x-rays jew forma ta' halqek, spejjez taghna.

DRITTIJIET TIEGHEK. Sehem f'dan l-istudju huwa min rajk. Ghandek id-dritt ittemm sehmek meta jidhirlek u tiddeciedi jekk il-materjal migbur jibqa' parti mill-istudju jew jigi distrutt. Ghandek ukoll id-dritt tiddeciedi liema proceduri tixtieq li isiru jew ma jsirux. Il-kunfidenzjalita' tigi mharsa f'kull publikazzjoni ta' *data* u ahna qatt ma naghtu informazzjoni genetika lil-hadd hlief fic-cirkostanzi specjali imsemmija izjed l-isfel, taht 'AFFARIJIET OHRA RELATATI LIL-PROPOSTA'. Tista' taghmel mistoqsjiet meta' trid waqt li tkun qed tiehu sehem.

RISKJU, INKONVENJENZA U SKUMDITA'. Ma hemm l-ebda riskju lis-sahha tieghek, lanqas ma hemm proceduri skomdi jekk tiehu sehem.

BENEFICCJI. Ma hemm l-ebda beneficcju materjali ghalik billi tiehu sehem. Izda ohrajn jistghu jibbenefikaw jekk fil-kas li jigi zviluppat test jew kura bhala risultat ta'dan l-istudju.

AFFARIJIET OHRA RELATATI LIL-PROPOSTA.

RICERKA FIL-FUTUR. L-iskopijiet taghna gew imsemmija fl'ewwel paragrafu ta' dan id-dokument. Ghas-skopijiet tar-ricerka ser niksbu xi celloli tieghek. Minkejja li jsiru testijiet mal-ewwel, ghandna mnejn nahznu ftit DNA go *freezer* ghal studji simili.

Id-DNA jigi mizmum ghal mhux izjed min sitt snin, sakemm ma jintuzax kollu. Apparti hekk, inzommu certu informazzjoni fir-*records* taghna.

Id-DNA mhux ser jintuza ghar-ricerka ohra minghajr permess specifiku tieghek.

Jekk taqbel li id-DNA tieghek jintuza il-quddiem, tkun tista tiddiskuti mac tabib jew kunsillier genetiku jekk hemm xi hsiebijiet jinkwetawk

INFORMAZZJONI MEDIKU MHUX MISTENNI. Filwaqt l-investigazzjoni hemm possibilta', avolja zghira, li niksbu informazzjoni mediku mhux mistenni dwar sahhtek jew l-ghamla genetika tieghek. Jekk dan l-informazzjoni jigi mahsub li jkollu x'jaqsam ma xi kura medika tieghek, ahna naghtuh lit-tabib tieghek biss u bil-permess tieghek dejjem.

Dan l-informazzjoni ma jistax jikkomprometti bniedem iehor. Jista jkun il-kas li jkun hemm bzonn ta' izjed informazzjoni jew li jinkisbu izjed celloli. F'dan il-kas naghmlu kuntatt mieghek permess tat-tabib tieghek.

RELAZZJONIJIET FAMILJARI. Filwaqt l-investigazzjoni hemm possibilta', avolja zghira, li niksbu informazzjoni dwar relazzjonijiet fil-familja e.z. kazijiet ta' adottazjoni jew ta paternita'. Dan it-tip ta informazzjoni ma jigi moghti lil-hadd, hlief fic-cirkostanzi straordinarju li dan l-informazzjoni jehtieg ghal-kura medika tal-individwi koncernati. F'dan il-kas l-informazzjoni jigi pprovdut lit-tabib li qed jaghti il-kura.

KUNFIDENZJALITA' Kull informazzjoni migbur jew miksub dwarek jew il-familja tieghek hu kkunsidrat kunfidenzjali u privat. Il-kotba u il-*files* relevanti jinzammu maqfulin go ufficcju fl-Universita ta' Malta. Il-kontroll ta' dan l-informazzjoni huwa f'idejk. Haddiehor, lanqas membri tal-familja tieghek, ma jista jkollhom access minghajr il-permess tieghek.

KOLLEZZJONI, RICERKA U HAZNA TA' MATERJAL GENETIKU. Ghandu ikun car li il-kampjun ta celloli tieghek jigi wzat ghar-ricerka biss. L-ebda ricerka jew testijiet ohra ma jsiru minghajr il-permess specifiku tieghek kif indikat taht. Id-DNA miksub mic-celloli tieghek jigi mahzun go *freezers* imsakkrin go bini imsakkar fl-Universita' ta' Malta. Dawn il-kampjuni jigu identifikati minn *codes* taghna. L-uniku bniedem li jkollu access ghal dawn il-*codes* ikun l-Investigatur Principali, Dr. Simon Camilleri. Dan jintrabat li johrog ir-rizultati taht dawn il-codes anonimizzati biss, sabiex ismek ma jinkixef lill-hadd

KUNTATTI.Dan l-istudju huwa taht il-harsien tal-Investigatur Prinicipali, Dr. Simon Camilleri. Tista tikkuntatjah fuq 23401876 jew email simon.camilleri@um.edu.mt b'xi mistoqsijiet relatati.

Appendix 2. Tables and Pedigrees

Table 22. Genetic conditions where failure of eruption is a feature.

Condition	**OMIM**	**Phenotype**	**Gene**	**Mode of Inheritance**	**Gene function/Pathogenesis**
Apert Syndrome	#101200	Ectopic teeth, delay/failure of eruption	*FGFR2*	Autosomal dominant	Preferentially expressed in osteogenesis
Cleidocranial dysplasia	#119600 Dental variant.0011	Ectopic teeth, delay/failure of eruption, supernumerary teeth	*RUNX2*	Autosomal dominant	Osteoblast-specific transcription factor, regulator of osteoblast differentiation
Primary Failure of eruption	#125350	Localised failure of eruption	*PTHR1*	Autosomal dominant	Activation of phospholipase and inhibition of adenylylcyclase through stimulation of inhibitory G proteins
Frontometaphyseal Dysplasia	#305620	Partial anodontia Delayed tooth eruption Retained deciduous teeth	*FLNA*	X linked recessive	Cytosketal component. Membrane ruffles.
Hypohydrotic ectodermal dysplasia	#305110	Ectopic teeth, delay/failure of eruption, oligodontia	*ED1*	X linked recessive	Regulatory role in both the NF-kappa-B and JNK pathways
	#224900		*EDAR/ EDARADD*	Autosomal recessive	
	#129490		*EDA3*	Autosomal dominant	
Osteogenesis	#166200	Delayed eruption,	*COL1α1,*	Autosomal dominant	Reduced amounts of

Imperfecta		dentinal dysplasia	*COL1α2*		normal collagen I.
Osteopetrosis	#166600	Failure of eruption, root deformation.	*CLCN7*	Autosomal dominant	Inactive Osteoclasts
	#259700		*TCIRG1*	Autosomal recessive	
Osteopathia Striata	%300373	Failure of eruption	*OSCS*	Autosomal/X linked dominant	Not known
	%166500				
Pyknodysostosis	#265800	Delayed eruption of deciduous teeth Delayed eruption of permanent teeth Hypodontia	*CATHEPSIN K*	Autosomal recessive	Major lysosomal protease in osteoclastic bone resorption
Acrodysostosis	%101800	Malocclusion Delayed tooth eruption Hypodontia	*Unknown*	Autosomal dominant	Unknown
Albright Hereditary Osteodystrophy	#103580	Delayed tooth eruption Enamel hypoplasia	*GNAS1*	Autosomal dominant	Regulator of Calcium Metabolism
Osteoglophonic Dysplasia	#166250	Failure of eruption	*FGFR1*	Autosomal dominant	Regulator of cellular responses
Oculo-Facio-Cardio-Dental Syndrome	#300166	Delayed eruption	*BCOR*	X linked dominant	Key transcriptional regulator during early embryogenesis
Wolf-Hirschhorn Syndrome	#194190	Retained deciduous teeth Ectopic teeth Delayed eruption	*WHSC1 deletion possibly with MSX1*	Isolated cases	Expressed ubiquitously in early development
Albright Hereditary Osteodystrophy	#103580	Delayed tooth eruption Enamel hypoplasia	*GNAS1*	Autosomal dominant	Regulator of hormone-sensitive adenylate cyclase
Chondroectodermal	#225500	Neonatal teeth	*EVC/EVC2*	Autosomal recessive	Essential role in

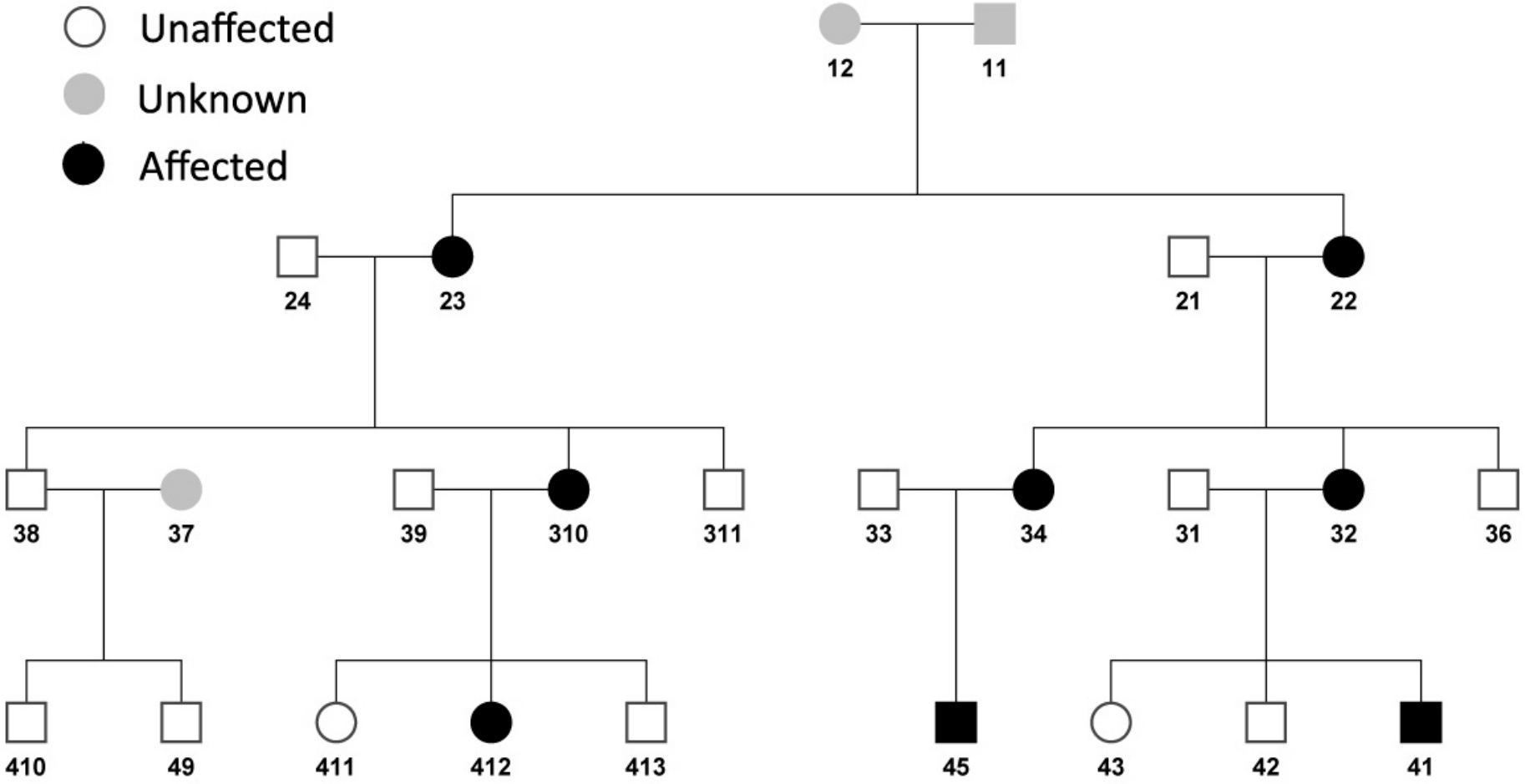
Dysplasia		Hypodontia Delayed eruption			skeletal development
Cornelia De Lange Syndrome	#122470	Widely spaced teeth Late-erupting teeth	*NIPBL*	Autosomal dominant Isolated cases	Chromatid cohesion and developmental regulation
Hunter Syndrome	+309900	Delayed tooth eruption Widely spaced teeth	*IDS*	X-linked dominant	Hydrolysis of 2-sulfate groups
Incontinentia Pigmenti	#308300	Hypodontia Delayed eruption Conical forms Accessory cusps	*NEMO*	X-linked dominant	Component of NFKB activation pathway
Focal Dermal Hypoplasia	%305600	Hypodontia Oligodontia Enamel hypoplasia Delayed eruption Malocclusion Notched incisors	*DHOF*	X-linked dominant	Not known
Pallister-Killian Syndrome	#601803	Delayed dental eruption	*Mosaicism for tetrasomy of 12p.*		Not known
Cockayne Syndrome	%216400 +133540	Delayed eruption of deciduous teeth Absent/hypoplastic teeth	*ERCC6*	Autosomal recessive	excision-repair cross-complementing group 8 gene
Mucopolysaccharidosis Type VI	+253200	Delayed eruption	*ARSB*	Autosomal recessive	Aryl sulphatase B deficiency
Hutchinson-Gilford Progeria Syndrome	#176670	Delayed and abnormal tooth eruption and morphology	*LMNA*	Autosomal dominant	Control of nuclear architecture and function

Table 23. The variants with a frequency of 1% or less in the 1000 genomes database which were common to both exomes.

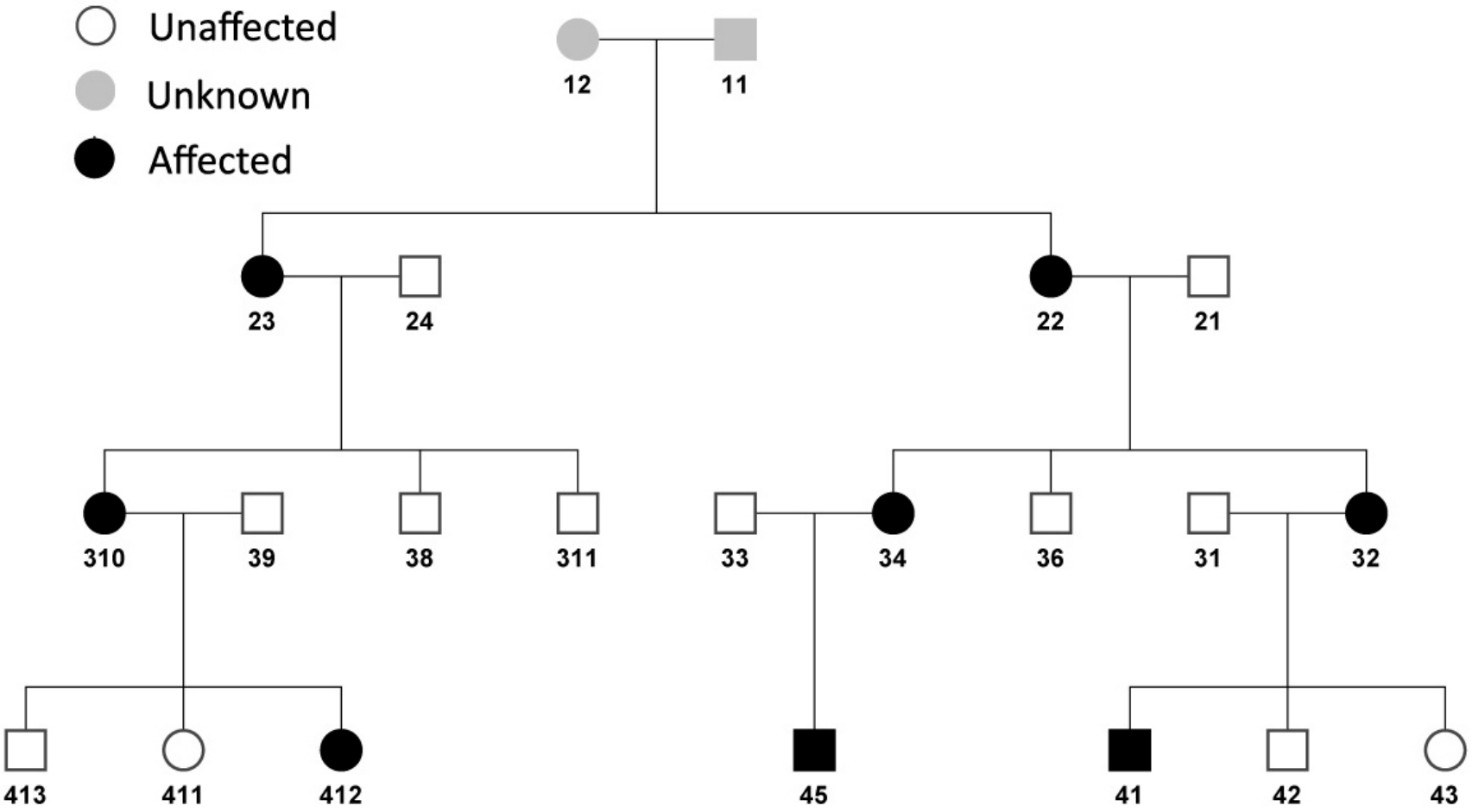
Chr	Position	Allele			details	SIFT score	dbSNP	1000g freq
		Ref	Alt					
2	61345243	G	A	exonic	KIAA1841:NM_032506:exon20: c.G2020A:p.A674T,	0.69	RS148264616	0.0005
3	37370566	A	T	exonic	GOLGA4:NM_001172713:exon17: c.A6240T:p.R2080S,	.	RS11924014	0.0055
3	38753732	A	T	exonic	SCN10A:NM_006514:exon22: c.T4009A:p.S1337T,	0.58	RS11711062	0.0014
4	110681527	C	T	exonic	CFI:NM_000204:exon6: c.G782A:p.G261D,	0.08	.	0.0009
6	26410148	T	C	splicing	BTN3A1:NM_194441:exon6: c.937+6T>C,	.	RS77721150	0.0064
6	31626080	C	T	splicing	C6orf47:	.	RS75629491	0.0077
6	32946119	G	C	exonic	BRD2:NM_001199456:exon9: c.G1654C:p.A552P,	0.22	RS55952113	0.0033
6	146007358	T	C	exonic	EPM2A:NM_005670:exon2: c.A376G:p.I126V,	1	RS150452237	0.0005
6	150464476	G	T	exonic	PPP1R14C:NM_030949:exon1: c.G148T:p.V50L,	0.45	RS200448500	0.004
6	160575837	G	A	exonic	SLC22A1:NM_003057:exon9: c.G1393A:p.G465R,	.	.	0.0082
8	68028252	C	G	exonic	CSPP1:NM_024790:exon11: c.C1376G:p.S459C,	.	RS146431326	0.0022
8	71646361	T	A	exonic	XKR9:NM_001011720:exon5: c.T824A:p.F275Y,	.	RS74941166	0.0041
8	86121636	T	C	exonic	E2F5:NM_001083588:exon6: c.T875C:p.I292T,	0.23	RS187526876	0.0023

8	87081697	C	G	exonic	PSKH2:NM_033126:exon1: c.G155C:p.R52P,	.	RS140592507	0.0047
8	87437445	A	C	splicing	WWP1:NM_007013:exon10: c.1062-7A>C	.	RS142212058	0.0031
8	106431420	A	G	exonic	ZFPM2:NM_012082:exon2: c.A89G:p.E30G,	0.05	RS121908601	0.0026
8	106814597	G	A	exonic	ZFPM2:NM_012082:exon8: c.G2287A:p.V763I,	0.11	RS117908591	0.0026
11	22284483	T	C	splicing	ANO5:NM_001142649:exon17: c.1798-9T>C,	.	.	0.0005
14	60027862	C	G	exonic	C14orf38:NM_001164399:exon7: c.G928C:p.D310H,	0.13	RS180781165	0.0007
14	64689913	G	A	exonic	SYNE2:NM_182913:exon8: c.G1141A:p.E381K,	0.48	RS150172232	0.0018
14	72054708	G	A	exonic	SIPA1L1:NM_015556:exon2: c.G119A:p.R40Q,	0.14	RS78621209	0.0069
18	31324766	G	A	exonic	ASXL3:NM_030632:exon12: c.G4954A:p.V1652M,	.	RS17746949	0.0085
19	39207742	G	A	exonic	ACTN4:NM_004924:exon10: c.G929A:p.R310Q,	.	RS112545413	0.0067
20	36560190	C	T	exonic	VSTM2L:NM_080607:exon2: c.C275T:p.A92V,	0.07	RS146022041	0.0045
22	24225971	G	A	exonic	SLC2A11:NM_001024939:exon9: c.G1006A:p.E336K,	.	RS79905160	0.0014

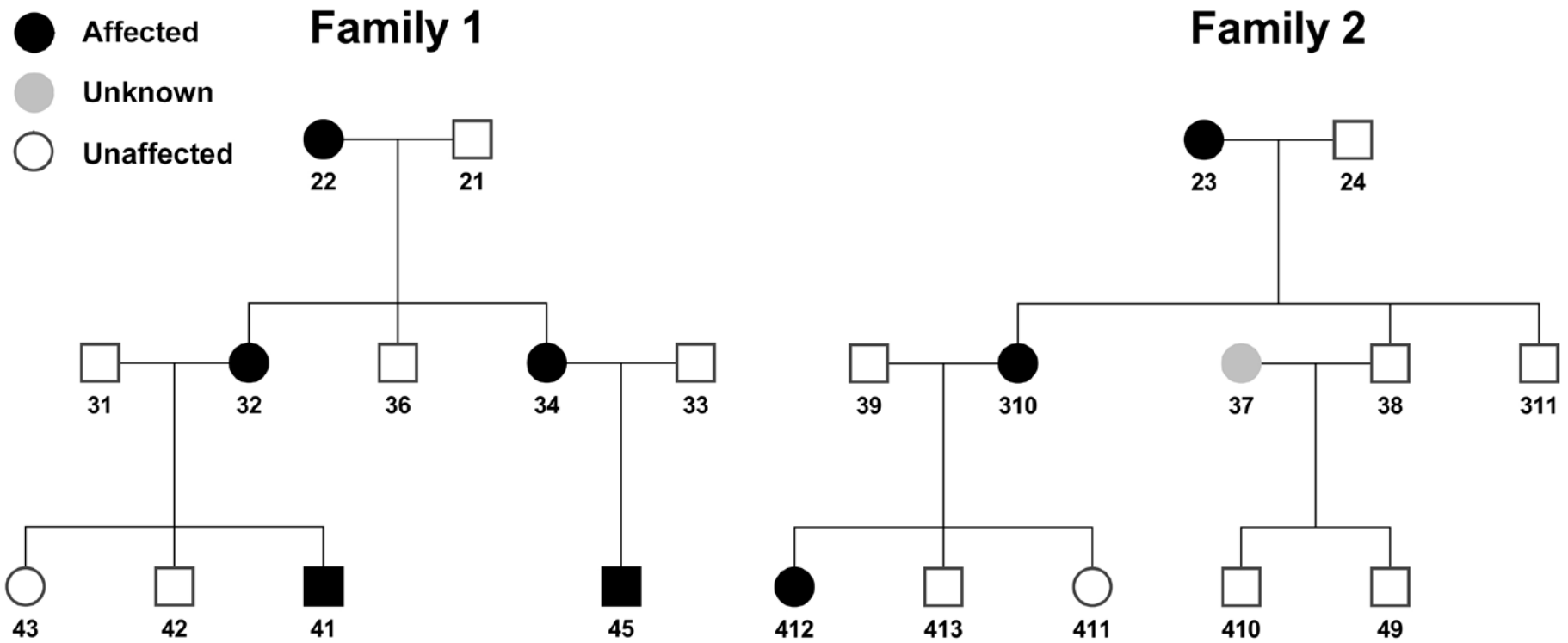
Pedigrees used in Linkage Analysis Study



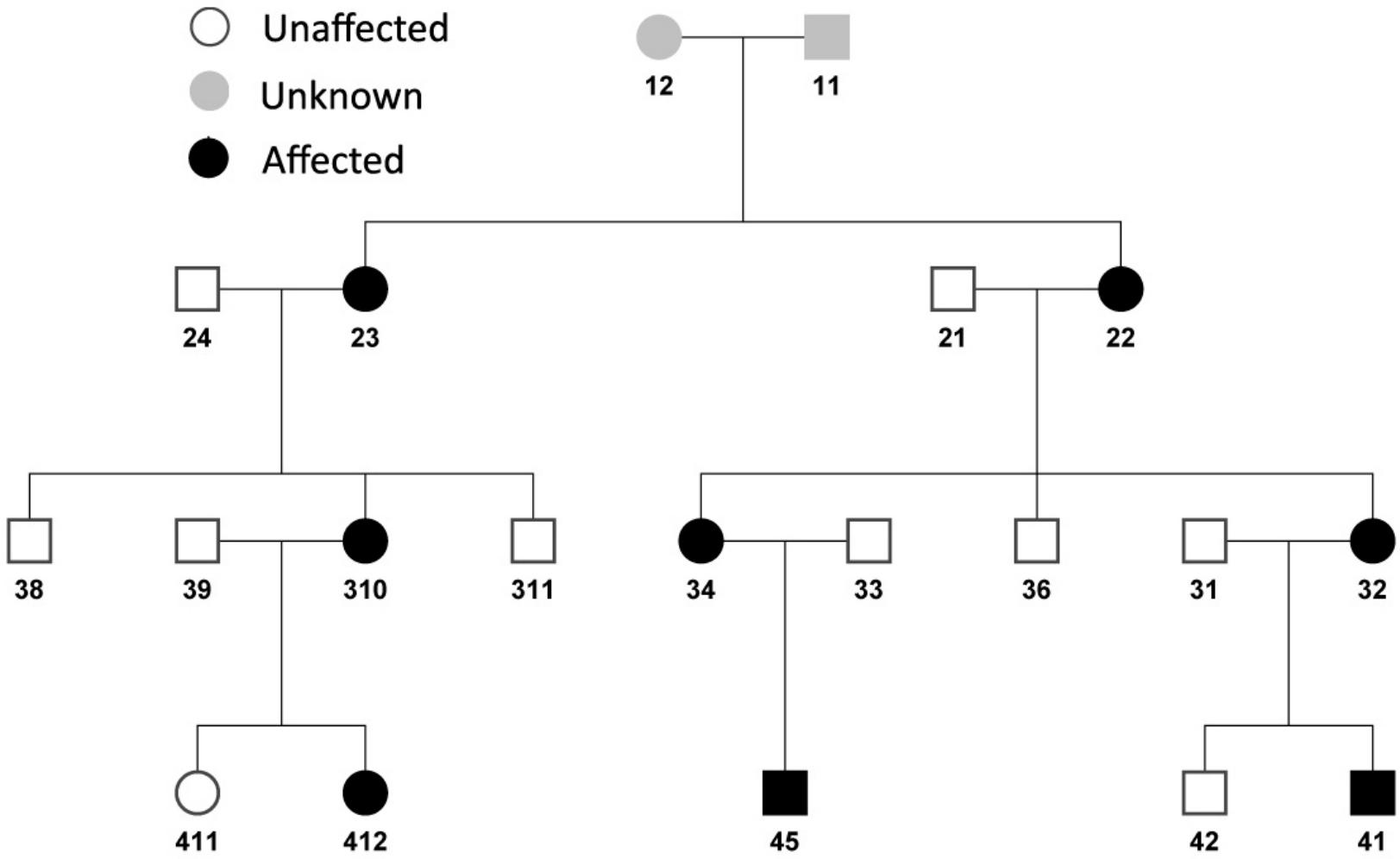
Pedigree V1, the full family tree.



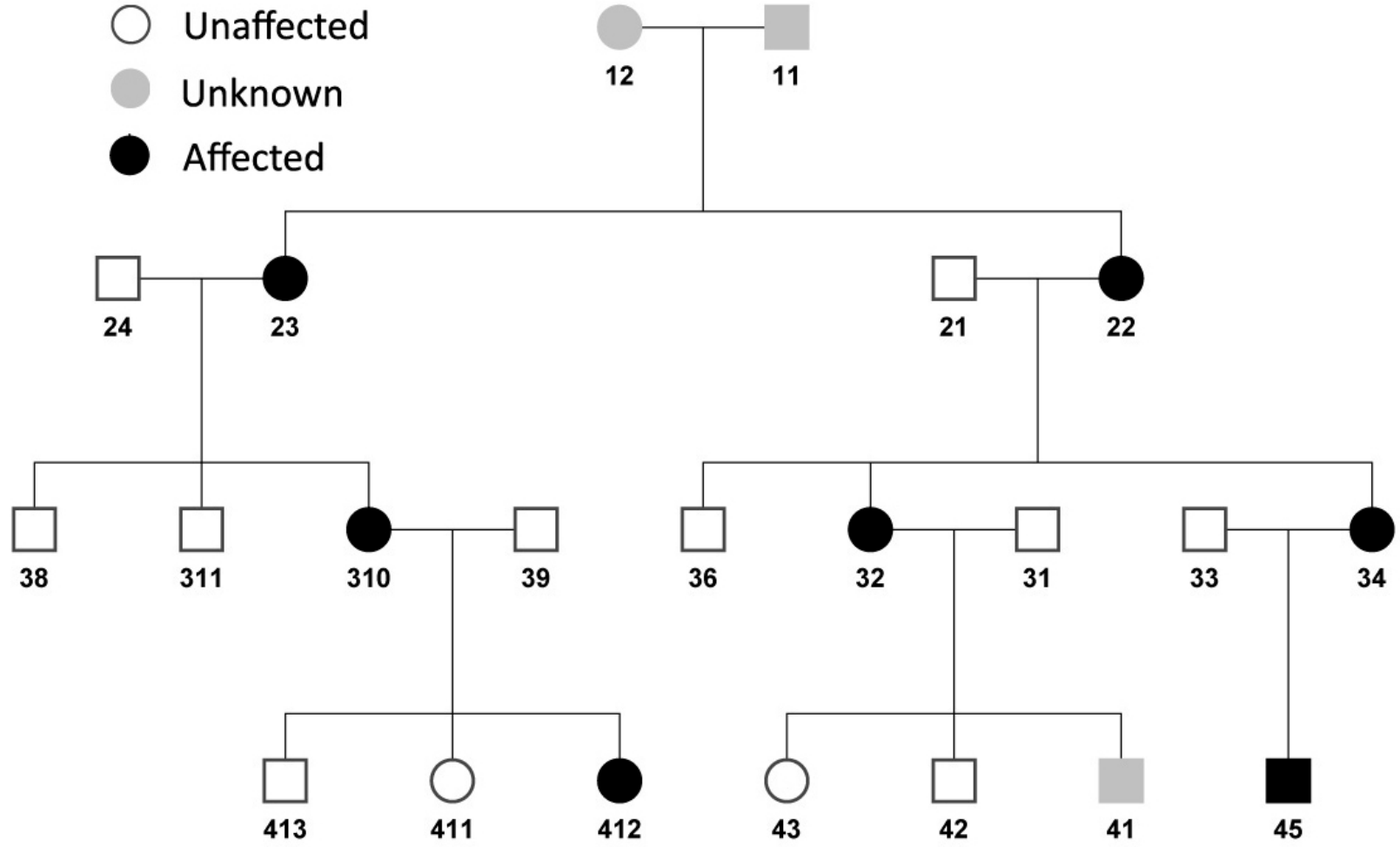
Pedigree V2, with uninformative family members removed.



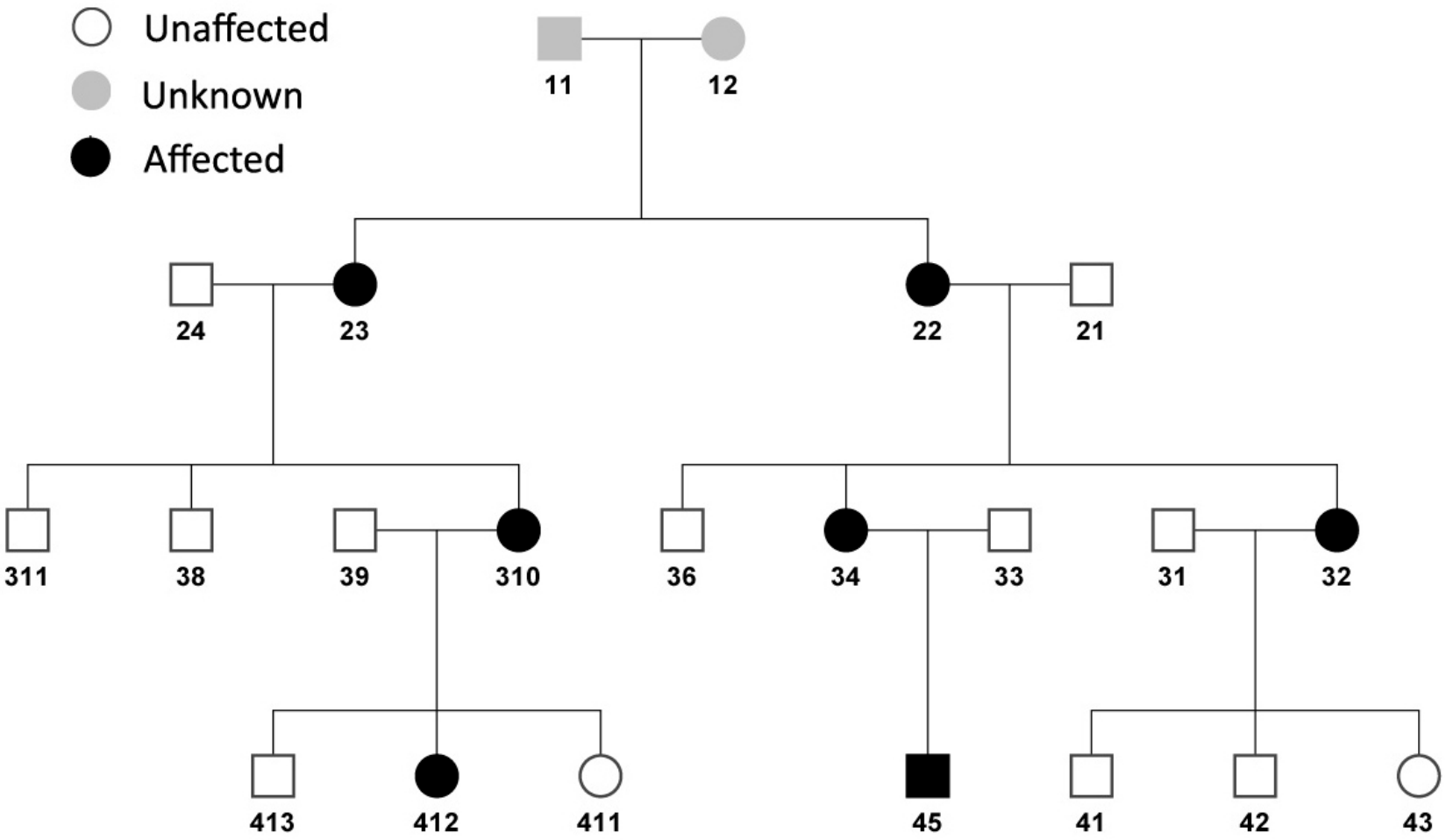
Pedigree V3, the full family, split into two.



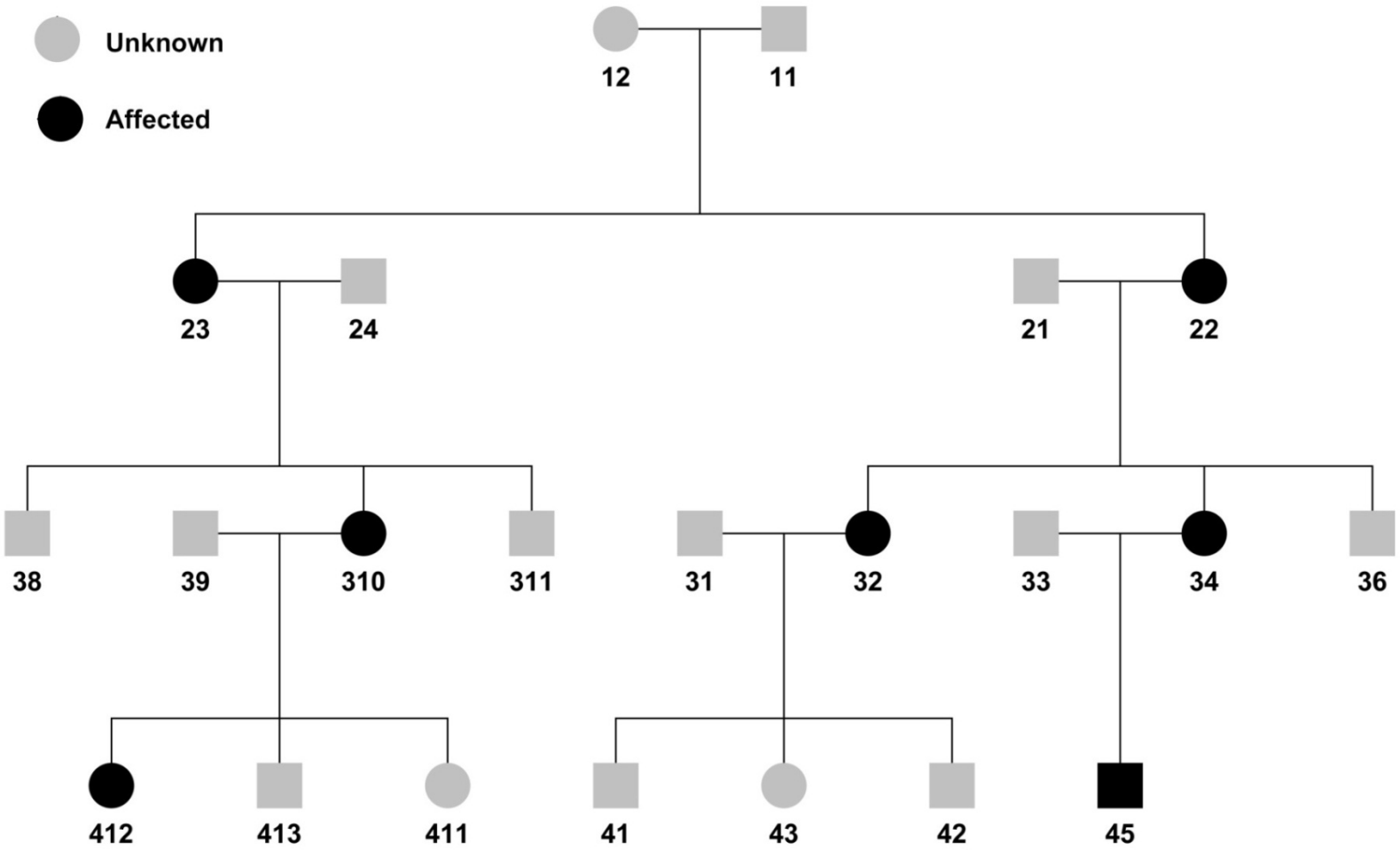
Pedigree V4, a further reduced version of V2, used for quality control.



Pedigree V6, identical to V2 but with #41 marked as 'status unknown'.

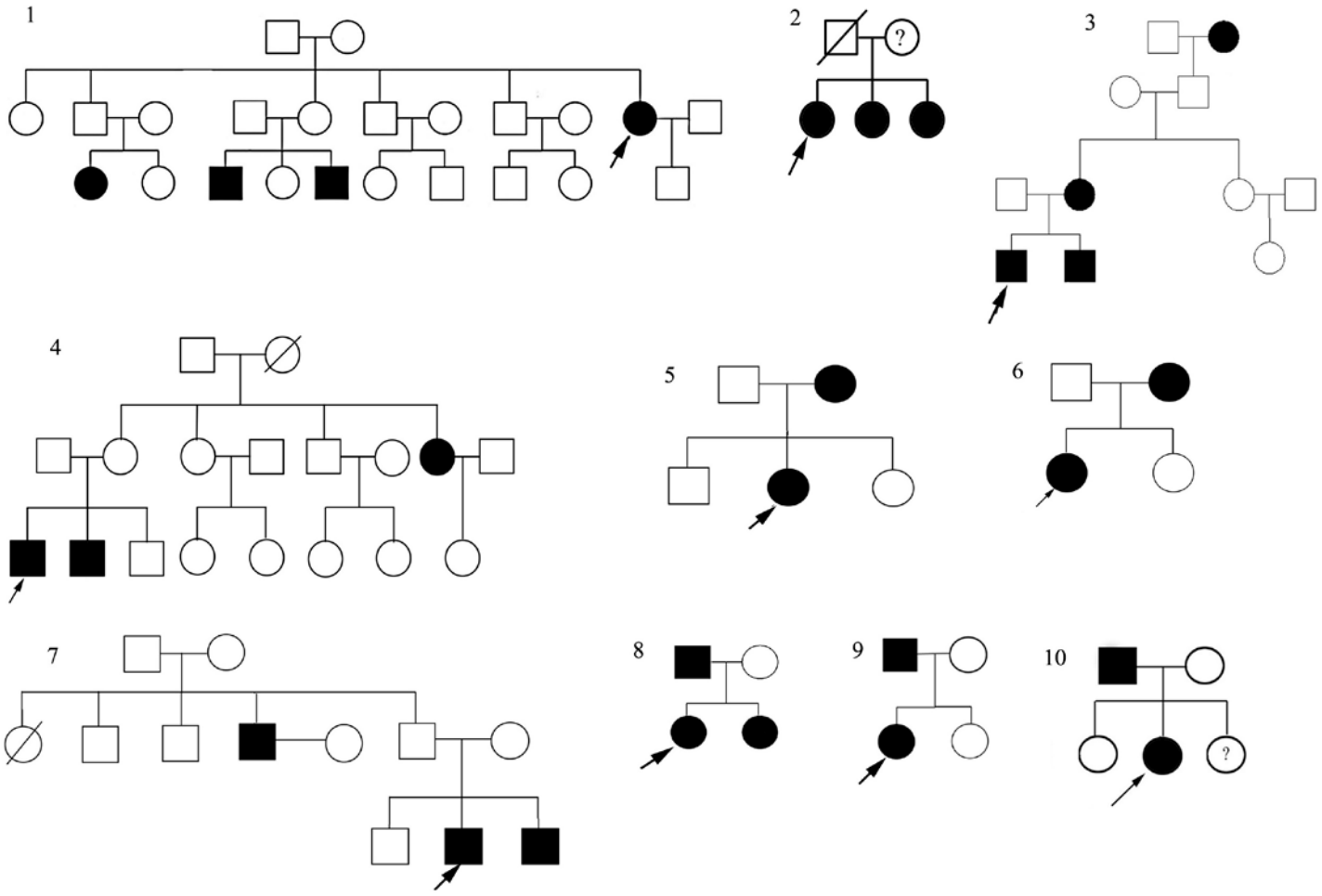


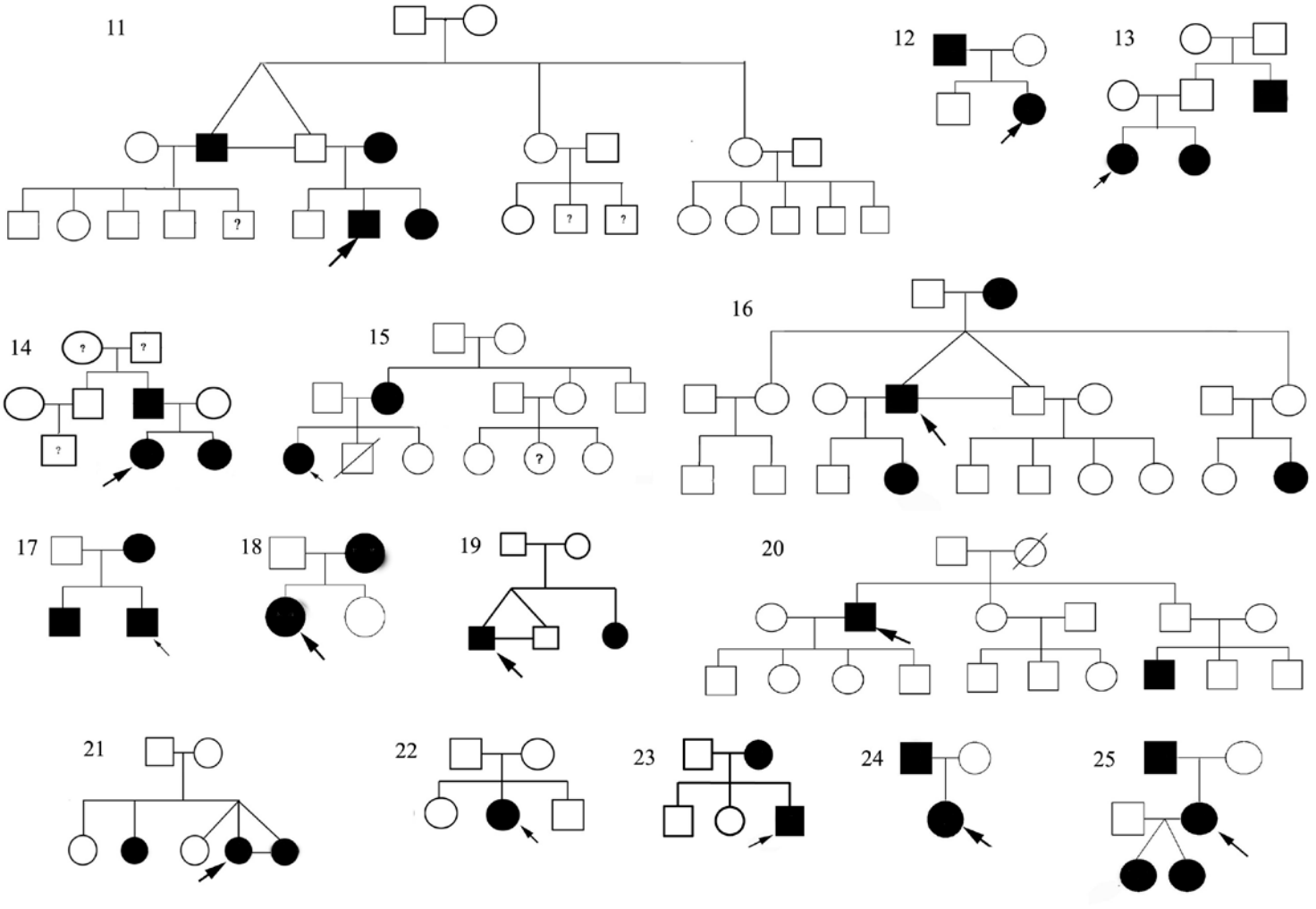
Pedigree V7, identical to V2 but with #41 marked as unaffected.

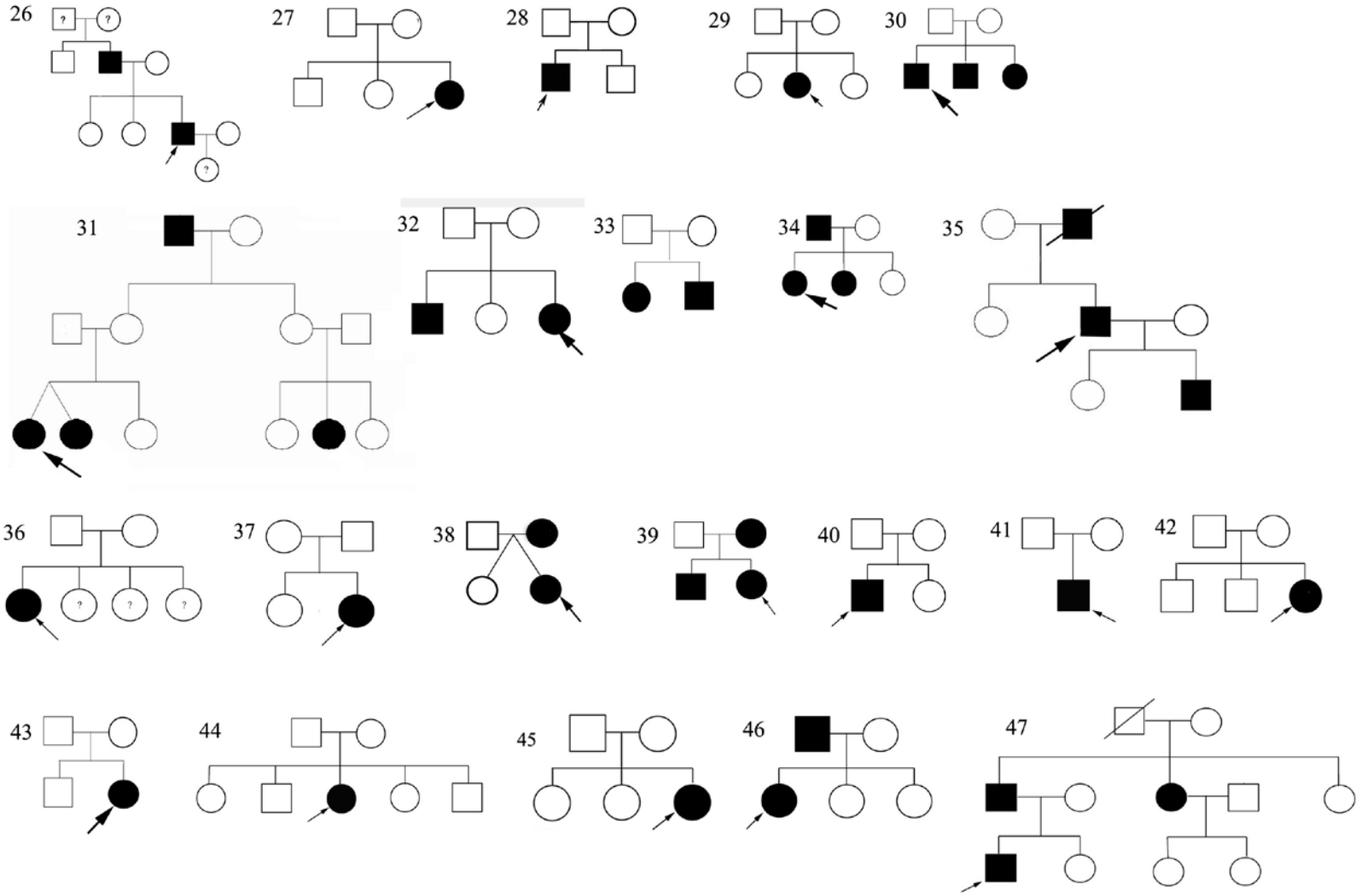


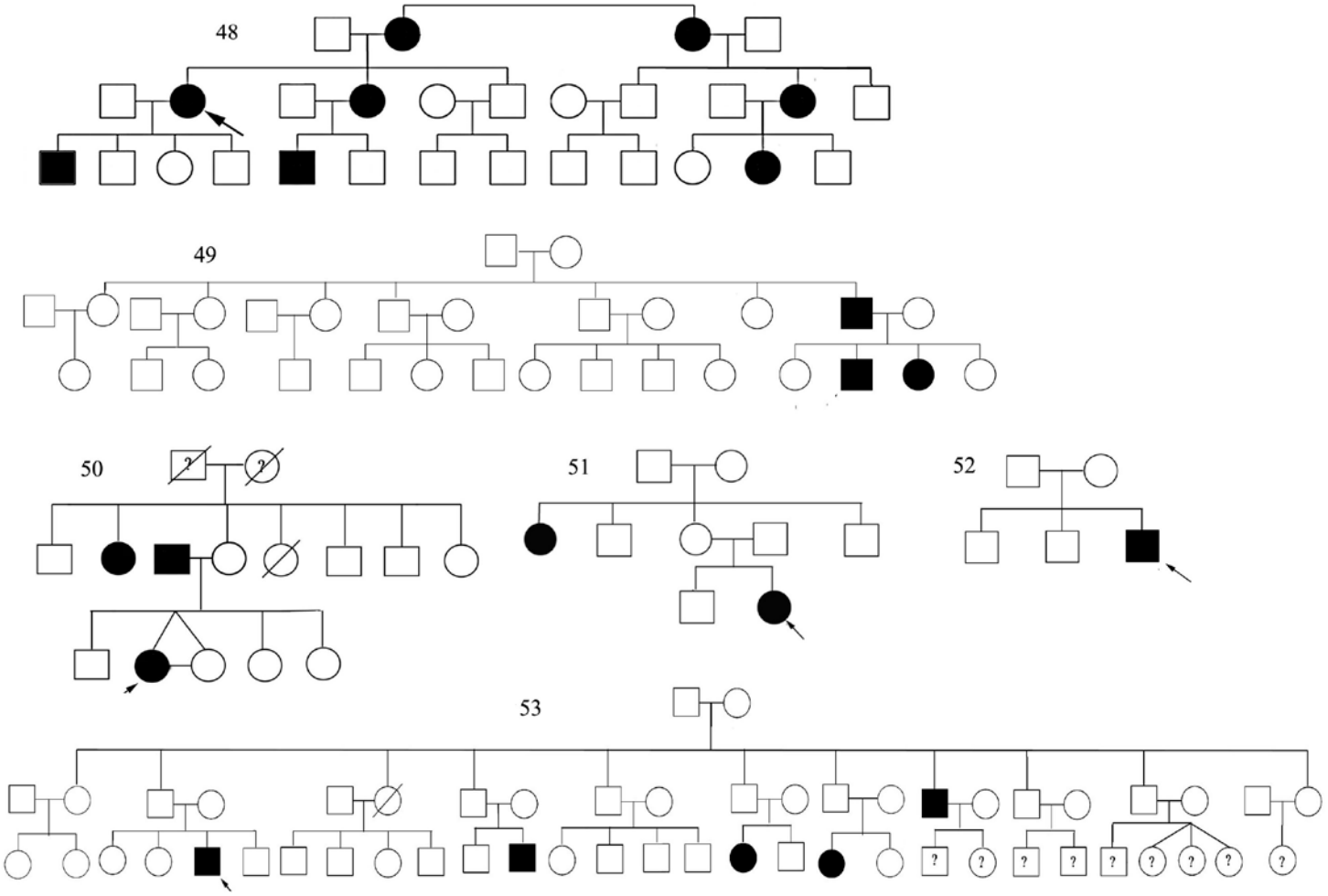
Pedigree V8, identical to V2 but with all clinically unaffected members and #41 classified as 'status unknown'.

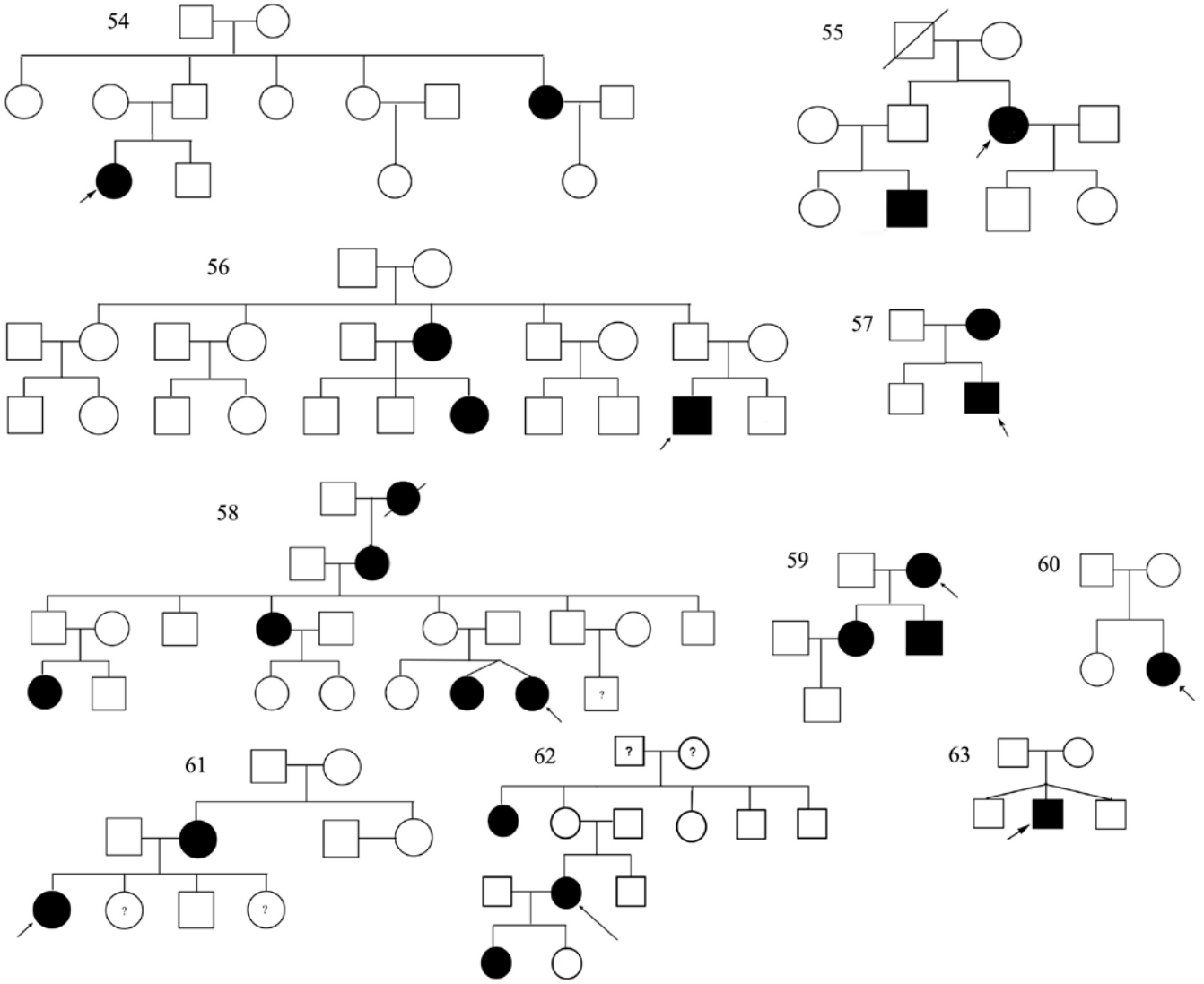
Pedigrees used in the segregation analysis study




11
12
13
14
15
16
17
18
19
20
21
22
23
24
25


26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47


48
49
50
51
52
53


54
55
56
57
58
59
60
61
62
63

Appendix 3. Publications

Abstract submitted to the British Society for Dental Research, March 2006.

Title of talk: **The Aetiology Of Ectopic Maxillary Canine Teeth**

Presenter: Simon Camilleri, Department of Orthodontics

Supervisors: Fraser McDonald, Christian Scerri

Objectives: To elucidate the mode of inheritance of ectopic canines.

Introduction: The aetiology of ectopic maxillary canines has been proposed to be genetic and is associated with incisor-premolar hypodontia as well as with various other anomalies.

The Maltese population has a high prevalence of ectopic teeth, especially ectopic canines, as compared to other populations. This has been ascribed to the 'Founder Effect' phenomenon, the population having grown from under 20,000 to over 400,000 in the past 500 years.

Methods: Probands with ectopic maxillary canines were identified during routine clinical examination. The inclusion criteria were a positive family history of ectopic canines, peg incisors or missing teeth in more than one generation or with an affected first cousin. Families with a member affected by a genetic syndrome were excluded.

Ectopic or missing teeth were identified by clinical examination or from existing records. Radiographic investigation was undertaken where necessary.

Twenty-two families were examined, comprising a total of 98 individuals. Pedigrees were constructed and the percentages of affected first-degree relatives determined.

Results: Analysis suggests autosomal dominant transmission. 12% of first degree relatives had ectopic canines, 4% had transposed canines and a further 12% exhibited hypodontia, in particular upper lateral and lower central incisors.

Penetrance is highly variable between families and there seems to be no clear pattern of augmentation or attenuation of symptoms.

One concordant pair of fraternal twins and one discordant pair of identical twins were included in the sample.

Conclusion: The genetic aetiology of ectopic canines is supported by this study, as is its close association with incisor-premolar hypodontia. Expressivity is also highly variable with other eruption anomalies such as maxillary and mandibular canine transposition forming part of the phenotype. The discordant identical twins suggest the possibility of epigenetic factors influencing eruption of teeth.

Abstract submitted to the American Society of Human Genetics Conference October 2006.

Title of talk: **The Aetiology Of Ectopic Maxillary Canine Teeth**

Presenter: Simon Camilleri, Department of Orthodontics

Supervisors: Fraser McDonald, Christian Scerri

Objectives: To elucidate the mode of inheritance of ectopic canines.

Introduction: The aetiology of ectopic maxillary canines has been proposed to be genetic and is associated with incisor-premolar hypodontia as well as with various other dental anomalies. The Maltese population has a high prevalence of ectopic teeth, especially ectopic canines, as compared to other populations. This has been ascribed to the 'Founder Effect', the population having grown from under 20,000 to over 400,000 in the past 500 years.

Methods: Thirty consecutive probands with a family history of ectopic canines were identified. 152 first, 51 second and 113 third degree relatives were contacted and their dental status ascertained by direct examination, anamnestic records, or written or telephone questionnaire. Pedigrees were constructed; the risk to the relatives was determined and plotted against the degree of relation.

Results: Analysis of the pedigrees suggests autosomal dominant transmission. 12% of first degree relatives had ectopic canines, 3.9% had transposed canines and a further 9% exhibited hypodontia, in particular upper lateral and lower central incisors. There is an appreciable relative risk in second and third degree relatives. The female to male ratio of the sample is 2.06, with no difference in the incidence of relatives of male or female probands. Penetrance

is highly variable as is expressivity, with wide variation in the number and severity of ectopicity of teeth.

Conclusion: The genetic aetiology of ectopic canines is supported by this study. The data points to the action of a single major gene with incomplete penetrance and variable expressivity.

Eur J Oral Sci 2006; 114: 361–373
Printed in Singapore. All rights reserved

European Journal of
Oral Sciences

Review

Runx2 and dental development

Simon Camilleri, Fraser McDonald

Department of Orthodontics, Dental Institute of Kings College London, London, UK

Camilleri S, McDonald F. Runx2 and dental development. Eur J Oral Sci 2006; 114: 361–373. © 2006 The Authors. Journal compilation © 2006 Eur J Oral Sci

The *Runx2* gene is a master transcription factor of bone and plays a role in all stages of bone formation. It is essential for the initial commitment of mesenchymal cells to the osteoblastic lineage and also controls the proliferation, differentiation, and maintenance of these cells. Control is complex, with involvement of a multitude of factors, thereby regulating the expression and activity of this gene both temporally and spatially. The use of multiple promoters and alternative splicing of exons further extends its diversity of actions. RUNX2 is also essential for the later stages of tooth formation, is intimately involved in the development of calcified tooth tissue, and exerts an influence on proliferation of the dental lamina. Furthermore, RUNX2 regulates the alveolar remodelling process essential for tooth eruption and may play a role in the maintenance of the periodontal ligament. In this article, the structure of *Runx2* is described. The control and function of the gene and its product are discussed, with special reference to developing tooth tissues, in an attempt to elucidate the role of this gene in the development of the teeth and supporting structures.

Department of Orthodontics, Floor 22 Guy's Tower, Dental Institute of Kings College London, St Thomas Street, London SE1 9RT, UK

Telefax: +44–20–71884415
E-mail: simon.camilleri@kcl.ac.uk

Key words: bone; odontogenesis; *Runx2*; tooth eruption

Accepted for publication July 2006

Mutations of *Runx2* [also known as Core-binding factor α1 (Cbfa1), PEBP2A1, and AML3) have been identified as being responsible for cleidocranial dysplasia (CCD) (Fig. 1) (1). This is an autosomal-dominant inherited disorder characterized by patent fontanelles, wide cranial sutures, frontal bossing, hypoplasia of the clavicles, short stature, ectopic and delayed eruption of teeth, supernumerary teeth, and other skeletal anomalies. Gene knockout experiments have produced similar skeletal phenotypes in mice (2).

RUNX2 is an osteoblast-specific transcription factor necessary for the differentiation of pluripotent mesenchymal cells to osteoblasts (3). The presence of RUNX2 in fully differentiated cells supports the concept that RUNX2 is also required in maintaining fully functional cells, at least in bone (3–5).

RUNX2 has also been identified as essential for tooth formation (4, 6). Dental abnormalities seen in CCD patients (7, 8) may be a direct result of RUNX2 dysfunction in tooth-forming cells. The dental abnormalities in CCD suggest an important role for RUNX2 during formation of the dentition.

The *Runx2* gene is 220 kb long (9) and contains eight exons (5, 10, 11). It belongs to the runt domain (RUNX) family of genes. These genes, named *Runx1*, *-2 and -3*, exhibit a high amino acid homology. Their protein products form a heterodimer with core-binding factor β (CBFβ) (12).

CBFβ is required for the efficient function of RUNX2 in skeletal development (13), allosterically enhancing DNA binding by RUNX proteins at the runt homology domain (RHD) (14). Furthermore, it plays an important role in stabilizing RUNX proteins against proteolytic degradation by the ubiquitin–proteasome system (15).

Runx2 and bone

Bone tissue consists of hydroxyapatite crystals and various kinds of extracellular matrix (ECM) proteins, including type I collagen, osteocalcin, osteonectin, osteopontin, bone sialoprotein (BSP), and proteoglycans (16, 17). These bone matrix proteins are secreted and deposited by polarized mature osteoblasts, which are aligned on the bone surface (18). The precise roles of matrix proteins in the formation of bone are not fully elucidated (19, 20). The formation of hydroxyapatite crystals is also regulated by osteoblasts.

The establishment of *Runx2* null mice clearly demonstrated that this transcription factor is essential for osteoblast differentiation, as these mice had no bone tissue, osteoblasts or osteoclasts, despite normal cartilaginous skeletal patterning. Chondrocyte maturation, however, is perturbed as a consequence (2, 21).

Runx2 is also active in mature osteoblasts. Mature mice, in whom active RUNX2 levels have been reduced, exhibit decreased expression of the genes encoding the main bone matrix proteins, BSP, osteocalcin, osteopontin and collagen type I (4). These genes are regulated via the RUNX2-binding sites in the proximal promoter segments of the respective genes (14).

RUNX2 has thus been shown to be essential for normal bone formation, with perturbation of bone formation if the levels are insufficient. Overproduction will also

Fig. 1. A 12-yr-old female patient with cleidocranial dysplasia. (A) Lateral skull radiograph showing a wormian bone in the temporal suture region and a hypoplastic maxilla. (B) Panoramic radiograph showing supernumerary and unerupted teeth, some of which are ectopic. Typically, the lower permanent central incisors and all four first permanent molars have erupted.

affect bone formation. Osteoblasts taken from non-syndromic synostosed sutures in children exhibited an increase in Runx2 expression, possibly explaining the enhanced proliferation and bone-forming ability of these cells (22). However, adult transgenic mice overexpressing Runx2 showed osteopenia with a decrease in bone mineral density. This was attributed to reduced osteoblast maturation, but also to enhanced receptor activator of nuclear factor kappa β ligand (RANKL) and matrix metalloproteinase-13 (MMP-13) production with enhanced osteoclastogenesis (23). Neonatal transgenic mice showed maturational blockage of osteoblasts, but did not show enhanced osteoclastogenesis, possibly because of the different ages of the experimental mice (24). The results of these experiments are consistent with evidence that *Runx2* is regulated, in part, through a negative feedback loop by activity of the RUNX2 protein on its own promoter, to control variations in gene expression and function during osteogenesis (25).

Runx2 expression decreases with age. This may be one possible explanation for impaired osteoblast function and reduced bone formation with aging (26).

RUNX2 is affected by a diversity of signaling pathways. Binding of ECM proteins to cell-surface integrins, mechanical loading, fibroblast growth factor 2 (FGF2), parathyroid hormone (PTH), and bone morphogenetic proteins (BMP), all influence RUNX2-dependent transcription. These post-translational changes act via the mitogen-activated protein kinase (MAPK) and protein kinase A and C (PKA, PKC) pathways, activating RUNX2 by phosphorylation (27–29) (Fig. 2). Further control of the RUNX2 molecule occurs through the counterplay of acetylation and ubiquitinization (29–31) (Fig. 3) The *Runx2* gene, in fact, plays a central role in co-ordinating multiple signaling pathways affecting osteoblast differentiation.

A species-specific difference in action may, however, exist. Osteoblast differentiation in rodents is associated with an increase in RUNX2 production. However, in human bone marrow stem cells, enhanced transactivation activity occurs with no change in Runx2 mRNA or protein levels. This activity has been ascribed to post-translational phosphorylation of key residues in the RUNX2 protein (32).

RUNX2 domains

RUNX2 binds to the core binding factor site, also known as the osteoblast-specific *cis*-acting element 2 (OSE2) (33), which is found in the promoter regions of all the major osteoblast-characteristic genes, such as osteocalcin, type I collagen, BSP, osteopontin, MMP-13, and alkaline phosphatase. Together with other factors, such as activator protein 1 (AP-1) and mothers against decapentaplegic homolog (SMADs), it controls their expression (34, 35). The RHD is responsible for the DNA-binding properties of RUNX2 (Fig. 4).

Three transactivation domains and one major repression domain have been identified in the RUNX2 protein (36). The first transactivation domain is located in the N-terminal 19 amino acids of the protein, while the second is located in the glutamine/alanine (Q/A) domain. In the latter, transactivation depends on a stretch of 29 glutamine residues. Deletion of the alanine stretch does not affect transactivation; however, expansion has a repressive function (1), and expansion may also play a role in nuclear localization (37).

The third activation domain is present in the N-terminal portion of the proline/serine/threonine (PST)-rich domain. A mutation in this region has been shown to cause a failure to interact with SMADs, reducing the response of osteoblasts to the Transforming growth factor-β/Bone morphogenetic protein (TGF-β/BMP) signaling pathway (38). This region has also been shown to interact with the co-activator molecule, p300, affecting expression of the osteocalcin gene (39). This action is independent of the acetyltransferase activity of p300, which protects RUNX2 from degradation by SMAD ubiquitin regulatory factor (SMURF)-mediated ubiquitination and also increases the transactivation potential of RUNX2 (31).

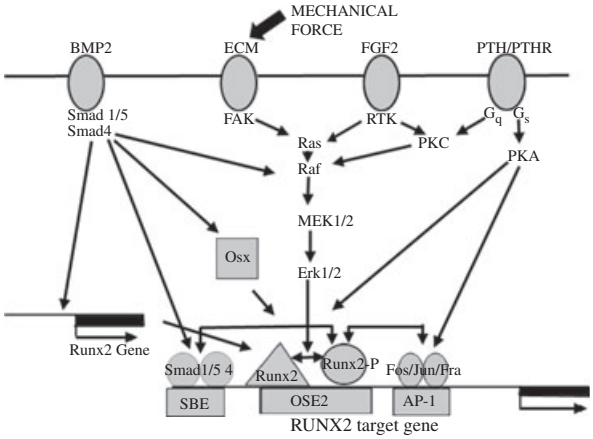


Fig. 2. Multiple signaling pathways converge on RUNX2 to initiate osteoblast differentiation. BMP2, Bone Morphogenic Protein 2; ECM, Extracellular Matrix; FGF2, Fibroblast Growth Factor 2; AP-1, Activator Protein 1; ERK, Extracellular signal related kinase; FAK, Focal adhesion kinase; Fos, Oncogene Fos; Fra, Fos related antigen; G, G alpha protein; Jun, Oncogene Jun; MEK1/2, Mitogen activated protein kinase/Extracellular signal related kinase1/2; OSE2, Osteoblast Specific Element 2; Osx, Osterix; PTH, Parathyroid hormone; PTHR, Parathyroid hormone related protein; Raf, Oncogene Raf; Ras, Retrovirus associated DNA sequences; RTK, Receptor Tyrosine Kinase; SBE, SMAD binding element; Smad, Mothers against decapentaplegic protein. Modified after Franceschi RT, Xiao G. J. Cell Biochem. Vol.88, 3 Pages: 446-454. Copyright © 2003 Wiley-Liss, Inc., A Wiley Company. Reprinted with permission of Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc.

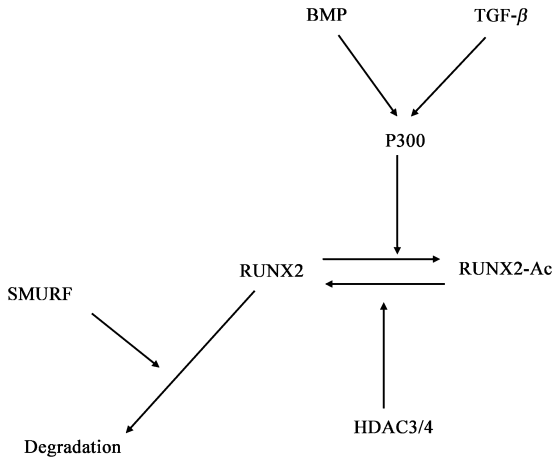


Fig. 3. Bone morphogenetic proteins (BMP) and transforming growth factor-β (TGF-β) act on P300 to acetylate (Ac) RUNX2, protecting it from degradation via the ubiquitin pathway. Histone deacetylases (HDACs) deacetylate RUNX2, accelerating its degradation. SMURF, SMAD ubiquitin regulatory factor.

The C-terminal part of the PST domain is a repression domain (36, 40). The terminal five amino acids (the VWRPY motif) are highly conserved and may bind the corepressor proteins of the transducin-like enhancer of split (TLE) or Groucho-related genes (Grg) family (41). The down-regulation of TLE/Grg expression during osteoblast differentiation is a potential mechanism for relief of Runx2 repression during cell differentiation (42). Other parts of the molecule have been shown to react with other corepressors, such as histone deacetylases (HDAC) (40), SIN3 (43), and yes-associated proteins (YAP) (44) (Table 1 and Fig. 4). A review of RUNX2 corepression was made by WESTENDORF (45).

RUNX2 is imported to the nucleus after transcription and must bind to specific regions of the nuclear matrix to effect transcriptional control (53), colocalizing with co-activators such as SMADs (35) and RNA polymerases, at nuclear sites that support RNA synthesis (54).

This function is effected by the nuclear matrix targeting signal (NMTS) region, a 38 amino acid segment situated between the RHD and PST domains (55). Point mutations in the NMTS region have been shown to affect the intranuclear localization of RUNX2, possibly affecting its interaction with target genes that are involved in osteolytic activity (38, 56).

The interaction of transcription factors with cellular signal transducers at particular points in the nuclear matrix may partly explain the tissue-specific action of the RUNX proteins and of transcription factors in general.

RUNX2 isoforms

RUNX2 has two major N-terminal isoforms, whose transcription from the *Runx2* gene is separately regulated by either a proximal promoter or a distal promoter (Fig. 5) (10). The product of the proximal promoter starts with the amino acid sequence, MRIPVD, and is known as Type I, PEBP2A or P56. That of the distal promoter product starts with the amino acid sequence, MASNSL, and is known as Type II, Til-1 or P57. OSF2 (Type III) is a splice variant of Type II, present only in the mouse. It starts with the sequence MLHSPH, and its function is similar to the Type II isoform. It is not essential for optimal transcriptional activation (57), and it may play a species-specific role in the regulation of transcription (10).

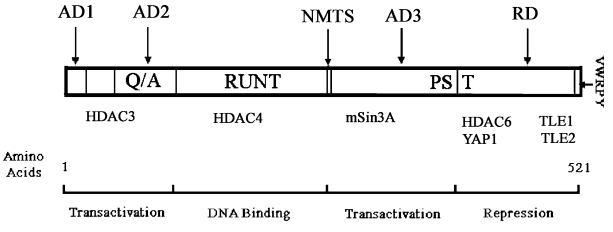


Fig. 4. Distribution of the different domains of the RUNX2 protein. The sites of interaction with various corepressor molecules are also shown. AD, transactivation domain; HDAC, histone deacetylase; SIN3A, SIN3A protein; NMTS, nuclear matrix targeting signal; PST, proline/serine/threonine rich domain; Q/A, glutamine/alanine-rich domain; RD, repression domain; RUNT, Runt homology domain; TLE, transducin-like enhancer of split; VRWPY, conserved repression signal; YAP, yes-associated protein.

Table 1

Known co-activators and corepressors of RUNX2. CBFB, core binding factor subunit B; GR65, Groucho related gene 5; HDAC3, histone decetylase3; HDAC5, histone decetylase 5; HDAC6, histone decetylase 6; MORF, monocytic leukemia zinc finger protein-related factor; pRb, retinoblastoma 1; Sin3A, Sin3 protein; TAZ, tafazzin; TLE/GROUCHO, transducin-like enhancer of split/Groucho related genes; TWIST, Twist protein; YAP1, yes associated protein 1

Protein	Reference	Class	Function
CBFB	Yoshida *et al.* (13)	Co-activator	Enhances DNA binding of RUNX2
GRG5	Wang *et al.* (46)	Co-activator	Dominant-negative inhibitor of larger TLE proteins
HDAC3	Schroeder *et al.* (47)	Corepressor	Blocks RUNX2 dependent transcription
HDAC4	Vega *et al.* (48)	Corepressor	Prevents DNA binding
HDAC6	Westendorf *et al.* (40)	Corepressor	Deacetylation of histones
MORF	Pelletier *et al.* (49)	Co-activator	Potentiates RUNX2 dependent transcription
MOZ	Pelletier *et al.* (49)	Co-activator	Potentiates RUNX2 dependent transcription
SIN3A	Fenrick *et al.* (43)	Corepressor	Alters subnuclear localization
p-300	Jeon *et al.* (31), Sierra *et al.* (39)	Co-activator	Acetylation of RUNX2/potentiates Vitamin D actions
pRb	Thomas *et al.* (50)	Co-activator	Potentiates RUNX2 dependent transactivation
TAZ	Cui *et al.* (51)	Co-activator	Potentiates RUNX2 dependent transactivation
TLE/GROUCHO	McLarren *et al.* (41) Javed *et al.* (42)	Corepressor	Colocalizes at subnuclear level
TWIST	Bialek *et al.* (52)	Corepressor	Prevents DNA binding of RUNX2
YAP1	Zaidi *et al.* (44)	Corepressor	Colocalizes at subnuclear level

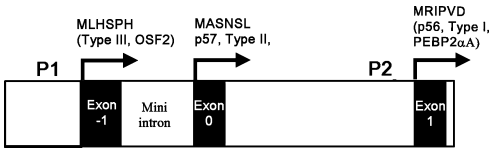


Fig. 5. Diagram of the *Runx2* promoter region. The P2 promoter drives expression of the Type I isoform and is associated with cell proliferation. The P1 promoter drives expression of the Type II variant, associated with commitment to a calcified tissue line. A mini-intron separates Exon 0 from exon −1. Transcription from this site produces the Type III variant, which has similar properties to Type II but is found only in mice. A, arginine; D, aspartic acid; H, histidine; I, isoleucine; L, leucine; M, methionine; N, asparagine; P, proline; R, arginine; S, serine; V, valine.

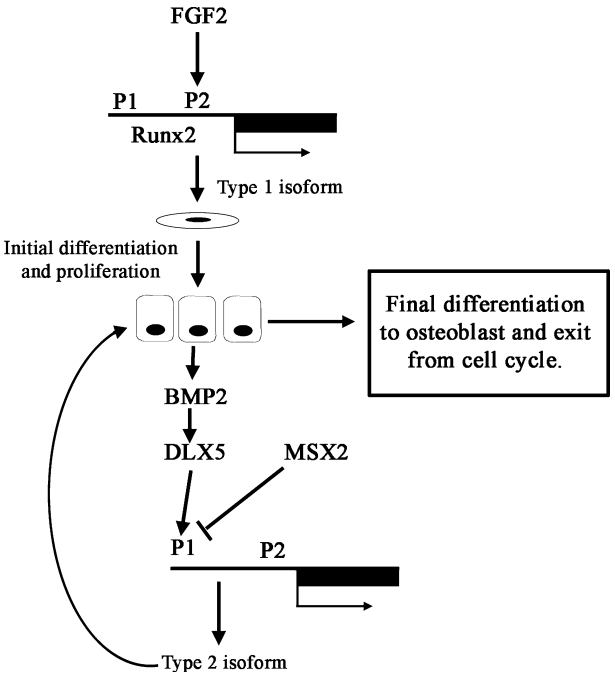


Fig. 6. Fibroblast growth factor 2 (FGF2) stimulates production of the Type 1 isoform from the P2 promoter to initiate cell commitment and proliferation. Bone morphogenetic protein 2 (BMP2), produced by these cells, acts via distal-less homeobox 5 (DLX5) on the distal P1 promoter to produce the Type 2 isoform, which encourages maturation of the osteoblast and exit from the cell cycle. The action of DLX5 is antagonized by muscle segment homeobox 2 (MSX2). Information derived from Bannerjee *et al.* (59), Choi *et al.* (64) and Lee *et al.* (65).

The presence of multiple isoforms of the *Runx2* gene product is consistent with other members of the *Runx* transcription family, including *Runx1* and *Runx3*, that exist as multiple isoforms with different transactivation potentials (58).

Type I and Type II isoforms differ functionally, but both are crucial in bone development. Type I is expressed ubiquitously in both non-osseous mesenchymal tissues and on osteoblast progenitor cells. It plays an important role in a wide range of cellular differentiation events and its expression is not affected by the differentiation status of the cell. In terms of intramembraneous bone formation, it probably is involved in the initial commitment steps and continues to exert its effects to the final differentiation of osteoblasts.

Type II expression is increased during osteoblast differentiation and is induced as a response to BMP2, suggesting that it is necessary for the maintenance of the osteoblast phenotype and that Type I may have a regulatory role (Fig. 6) (59).

In contrast to the findings of Ducy *et al.* (3) and Xiao *et al.* (10), Harada *et al.* (60) found all three isoforms to be present in the bones of adult mice and concluded that all three are involved in stimulating osteoblast differentiation, exerting different functions in the process of differentiation depending on the cell type and the

downstream gene promoter. Because production of the Type II isoform is restricted to later events of cell differentiation, being found only in pre-osteoblasts and osteoblasts, its contribution is probably more specific in the final step into osteoblast differentiation and is necessary for osteoblast maturation and skeletogenesis (61, 62). However, the expression pattern and function of both isoforms is not very different in the later stages of osteoblast differentiation (63).

The expression of Type I Runx2 is stimulated by FGF2. This commits precursor cells to the osteoblast lineage and permits cellular proliferation. Runx2 Type I stimulates production of BMP2 and this, in turn, affects *Runx2*, its effect being mediated by distal-less homeobox 5 (DLX5) (64). Lee *et al*. (65) found that DLX5 specifically transactivates the *Runx2* distal (P1) promoter in committed osteoblasts and that its action is antagonized by MSX2 (Fig. 6).

The pattern of isoform expression may be different in tooth tissue, with a high expression of Type I and weak expression of Types II and III in the ameloblast and odontoblast layers of neonatal mouse incisors (66). However, Types II/III have been shown to be indispensable for tooth germ development past the bud stage and they play a role in the differentiation of ameloblasts and odontoblasts, which is necessary for transcription of dentine matrix protein 1 (DMP1), dentine sialoprotein (DSPP), amelogenin (AMGN), and ameloblastin (AMBN) in cultured mouse mandibles (67).

Dual translational control of *Runx2*, via both the internal ribosomal entry site (IRES) and cap-dependent mechanisms, exists for the P1 and P2 promoters. This provides another level of control of RUNX2 isoforms and may be a means to fine-tune Runx2 expression across a wide range of cellular conditions where the amount of *Runx2* gene products may be an important determinant of their biological effects (68).

Further isoforms of RUNX2 exist as a result of alternative splicing. This also is consistent with findings for other RUNX factors and may explain the different spatiotemporal patterns of expression of the gene (10, 58, 69, 70). The timing of expression of different RUNX2 isoforms, in conjunction with other transcription factors and depending on the level of differentiation of the cell, may well serve as a method of control of bone and tooth formation.

Runx2 and osteoclastogenesis

Overexpression of Runx2 increases osteoblast number but inhibits their terminal maturation, resulting in accumulation of less mature osteoblasts and consequent osteopenia (24). Osteoclastogenesis is increased, possibly by the increased production of RANKL and MMP-13 (two factors involved in bone formation–resorption coupling) by the immature osteoblasts (23).

Osteoprogenitor cells have greater potential to support osteoclast development than more differentiated cells (71). Osteoclastogenesis is strongly induced by undifferentiated stromal marrow cells, which produce high RANKL levels. As maturation proceeds, RANKL levels drop and those of its antagonist, osteprotegerin (OPG), rise (72). The stimulus to osteoclastogenesis is reduced accordingly and the differentiated osteoblasts proceed to the formative phase of bone formation.

The precise role of RUNX2 in osteoclast regulation is controversial. Opg is strongly expressed in *Runx2*$^{-/-}$ calvarial cell lines, whereas Rankl is not (73, 74). However, RUNX2-binding elements are present in the promoter region of the *Opg* gene. RUNX2 increases the activity of the Opg promoter, suggesting that RUNX2 regulates osteoclastogenesis by inducing the expression of Opg (75). TGF-β inhibits bone resorption by induction of Opg and its effects are mediated by RUNX2 and SMADs (76).

A RUNX2-binding site is present on the mouse and human *Rankl* promoter (77), but RUNX2 failed to stimulate the transcriptional activity of the promoter region of the *Rankl* gene, showing that that RUNX2 possibly does not regulate *Rankl* in the same manner as other known targets (78). However, RUNX2 induced Rankl expression and suppressed Opg expression in the presence of $1,25(OH)_2D_3$ (74). On the other hand, forced production of soluble RANKL was found to be insufficient for the complete rescue of osteoclast differentiation in *Runx2*$^{-/-}$ mice, suggesting the presence of another requirement for osteoclast differentiation. Also, treatment of RUNX2-deficient calvarial cells with $1,25(OH)_2D_3$, affected both Rankl and Opg expression and induced osteoclastogenesis, showing that expression of Rankl and Opg, and initiation of osteoclastogenesis, may be induced via alternative pathways (79).

Colony-stimulating factor 1 (CSF-1) and RANKL are widespread; however, osteoclasts are confined to calcified tissue. The addition of CSF-1 and RANKL to serum-free cell cultures produced no osteoclasts (80). This suggests the existence of other factors or ECM proteins, possibly induced by RUNX2 in osteoblasts, which are required for osteoclast differentiation.

RUNX2 in odontogenesis

In the developing mammalian tooth, the cranial neural crest-derived dental mesenchyme consists of the dental papilla and dental sac. The dental papilla gives rise to dental pulp and odontoblasts; the dental sac gives rise to the periodontium, including the osteoblasts that contribute to the alveolar process. The alveolar process is a specialized intramembranously formed bone that provides the primary support structure for the dentition.

The expression of Runx2 in both the dental papilla and dental sac suggests a potential involvement of this gene in the differentiation of odontoblasts and osteoblasts lining bone in the periodontal space (6, 81). Mice deficient in RUNX2 exhibit an arrest of molar tooth development at the early cap stage, suggesting a requirement for RUNX2 in the progression of tooth development from the cap stage to the bell stage (6).

In mouse embryonic mandibular first molar tooth germs, Runx2 expression is initiated by FGF produced by the odontogenic epithelium, shortly after commencement of epithelial thickening, and is followed by expression of RANKL in the early alveolar bone ossification centers and that of its receptor, receptor activator of nuclear factor kappa B (RANK) and OPG in tooth bud epithelium and mesenchyme. Thus, *Runx2* is not involved in the initiation of tooth formation, but is intimately involved in regulating the expression of mesenchymal molecules that act reciprocally on the epithelium to control the histo- and morpho-differentiation of the enamel organ (6, 82, 83).

Runx2$^{-/-}$ mouse molars show arrested development at the bud stage, whereas incisors, which develop earlier, progress to the bell stage and show dentine formation, although odontoblasts are abnormal and no enamel is formed (6, 84).

Enamel knot marker genes, including cyclin-dependent kinase inhibitor 1A (*p21*), *Fgf4*, ectodysplasin A receptor (*Edar*) and *Bmp4*, are down-regulated in *Runx2*$^{-/-}$ lower molars, but are expressed normally in the upper molars. Sonic Hedgehog (*Shh*) is completely absent in *Runx2*$^{-/-}$ lower molars, while weak signals remain at the tip of the tooth bud in the upper molars (83).

Lower molars are more severely affected than upper molars in *Runx2*$^{-/-}$ mice, and incisors are less affected than molars. Hence, RUNX2 may have different downstream target genes in different regions of the jaws. Similar regional differences in molecular regulation is evident in relation to other genes, notably *Dlx* (85) and *Activin*βA (86). The different origins of the neural crest cells populating the maxillary and mandibular primordia may explain their different behavior (87). However, the developmental profiles of Runx2 expression in odontoblasts and osteoblasts, both derived from mesenchyme, is also different (6), suggesting that the gene may also be differentially regulated in these cells.

Runx2 is thus essential for tooth development up to the bell stage, being necessary for the formation of the enamel knot, which controls growth and folding of the enamel organ epithelium. Whether Runx2 is essential for the later stages of tooth development is still unknown, as *Runx2*$^{-/-}$ mice do not survive beyond birth.

Prior to crown development, Type II Runx2 is strongly expressed in dental papilla mesenchyme, which gives rise to the pulpal cells and odontoblasts. Type II Runx2 is markedly down-regulated at the bell stage in the dental papilla, after morphogenesis is complete. Expression continues, albeit at a lower level, in the cells of the dental papilla, particularly near the apical portion, as well as in the odontoblasts lining the pulp chamber. Expression is evident throughout the further development of the tooth and at all stages of root formation, including formation of the periodontium. Cementoblasts, cementocytes, periosteal tissue, osteoblasts, and osteocytes all showed expression of Type II Runx2. No expression was found in osteoclasts. The forming periodontium contains a decreasing gradient of transcripts and immunostaining from crown to the root tip (88) (Figs 7, 8).

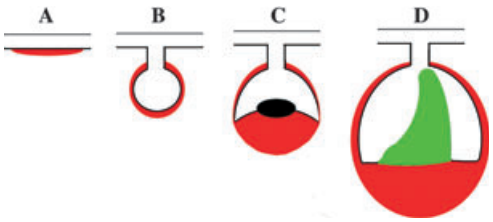


Fig. 7. Patterns of Runx2 expression. (A) Initiation stage. Expression is induced in the mesenchyme by the odontogenic epithelium. (B) Bud stage. Expression is maintained around the ingrowing dental epithelium. (C) Cap stage. Expression is sustained in the mesenchyme and the dental follicle. RUNX2 is necessary for Sonic Hedgehog (Shh) expression and for formation of the enamel knot. (D) Bell stage. In this stage, expression is down-regulated in the dental papilla but is maintained in the dental follicle and surrounding mesenchyme.

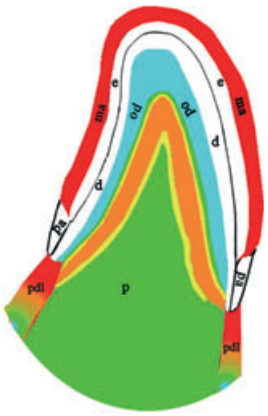


Fig. 8. Runx2 mRNA expression in the secretory stage. Pre-ameloblasts show no expression; however, this is strongly up-regulated in maturational phase cells. There is diffuse expression throughout the dental papilla, with up-regulation in pre-odontoblasts. Differentiated odontoblasts, however, show no expression. The periodontal ligament shows strong expression as do cementoblasts and cementocytes. d, dentine; e, enamel; od, odontoblasts; ma, maturational amleoblasts; p, dental papilla; pa, pre-amleoblasts; pdl, periodontal ligament.

Ameloblasts

Runx2 expression controls downstream factors acting on the development of the enamel organ epithelium. The importance of *Runx2* in amelogenesis is evidenced by the lack of enamel in the incisor tooth germs of *Runx2*$^{-/-}$ mice. RUNX2 is also present in late secretory- and maturation-stage ameloblasts (6).

AMBN is an extracellular matrix protein that may play a role in enamel crystal formation in the developing dentition. It is used as an ameloblast-specific gene marker. The murine *Ambn* promoter contains two RUNX2-binding sites. RUNX2 interacts with functionally important regions of the *Ambn* promoter, and mutations of the *Ambn* promoter's RUNX2-binding sites diminish promoter activity. This indicates that RUNX2 possesses an important function in transcription of the *Ambn* gene (89). Runx2 Type II mRNA remains strongly expressed in both immature and mature ameloblasts (81). Thus,

Runx2 is involved in the early stages of enamel organ formation, and tooth morphogenesis and may also play a direct role in formation of tooth enamel. Histological investigation of extracted permanent teeth in CCD show evidence of hypoplasia (90, 91). This does not appear to be the case for deciduous teeth (92).

Odontoblasts

Osteoblasts and odontoblasts share several similarities, including the expression of similar genes. Indeed, the main non-collagenous components of the odontoblastic extracellular matrix (DSPP or DMP1) are also present in other tissues, such as osteoblasts (93) and periodontium (94), albeit at much lower levels.

Multiple RUNX2-binding sites have been identified in the regulatory elements of the mouse *Dspp* gene (66). RUNX2 increased Dspp expression in immature odontoblasts, but down-regulated expression in more mature cells, showing that the effect of RUNX2 is dependent on the state of differentiation of the target cell (95).

RUNX2 is also involved in the regulation of DMP1 in osteoblasts, although it is not essential for DMP1 expression in odontoblasts, indicating the involvement of other unidentified odontoblast-specific transcription factors or co-activators (96, 97). Jiang *et al*. (81) found Type III Runx2 expression in immature odontoblasts at all stages, including cells in the dental papilla, confirming their potential to differentiate to odontoblasts.

Growth and differentiation in $Runx2^{+/-}$ human primary pulp cells are different to $Runx2^{+/+}$ cells as a result of variations in gene expression patterns and signaling (98). Therefore, the effects of haploinsufficency may well influence the differentiation of odontoblasts from these cells. Runx2 is up-regulated in early odontoblasts (99), showing that levels of RUNX2 are necessary at this stage. However, unlike in osteoblasts, Runx2 expression is remarkably low or undetectable in differentiated odontoblasts (88, 100).

This down-regulation of expression in newly differentiated and functional odontoblasts suggests that *Runx2* plays an essential and stage-specific role in the lineage determination and terminal differentiation of odontoblasts from dental papilla mesenchyme (101). It also highlights the different effects of this gene in different tissues.

Cementum

The origins of cementum-producing cells, and whether they share common precursors with osteoblasts, are unclear (102). Most cells embedded in the cellular cementum express Runx2 mRNA and RUNX2 protein to various degrees (88). As cementoblasts are mineralizing cells, this is not a surprising finding.

The result of *Runx2* haploinsufficiency on cemental tissue seems to vary between species. Both acellular and cellular cementum formation is defective in permanent teeth in CCD subjects. This does not seem to be the case in the deciduous teeth of CCD subjects (7, 103).

Reports on cementum formation in $Runx2^{+/-}$ mice are inconsistent (104, 105). Chung *et al*. (105) explained the similarity between mutant and wild-type mice in the study of Zou *et al*. (104) by the mice being too young to have developed their full root length. However, cementum formation in primary teeth in subjects with CCD seems to be normal (92). If the mouse dentition does represent the human primary dentition, gross abnormalities are not expected to be found in heterozygous mice.

Reports of root morphology in CCD literature varies from excessive root length (106), deformation (8,107), and 'spiky' (108).

Periodontal ligament

Type II Runx2 is also expressed in periodontal ligament (PDL) fibroblasts, although BSP, a marker of osteoblast differentiation and biomineralization, is not (81). The action of RUNX2 seems to be suppressed by a mechanism designed to maintain PDL width (109). The factor S100A4 has been implicated (110).

However, these cells maintain the potential to differentiate to osteoblasts under certain conditions, such as mechanical stress (111). Deformation of PDL osteoblasts increases Runx2 expression, protein levels, and also its DNA-binding activity, the latter possibly being caused by activation of the ERK (extracellular signal related kinase) MAPK pathway (34), although the bone response to stress is similar in both heterozygous and wild-type mice. Orthodontic tooth movement is not affected in $Runx2^{+/-}$ mice (105); one functioning gene seems to be sufficient to produce an adequate bone response.

Tooth maturation and eruption

Runx2 controls the maturation of both osteoblasts and odontoblasts. Therefore, a delay in tooth maturation is expected in RUNX2-deficient tissues. This is reflected in the clinical situation, where the dental maturation of CCD subjects is retarded by as much as 4 years when compared with unaffected subjects (8, 106, 112).

Zou *et al*. (104) found no difference in dental development or eruption timing between heterozygous knockout and wild-type mice. One functioning allele seems to be sufficient for normal dental development in the mouse. Yoda *et al*. (113) reached a similar conclusion as regards tooth development, but found a significant difference in tooth eruption times. This was explained as being caused by a time-specific lack of osteoclastic response, suggesting that heterozygous mice cannot recruit sufficient osteoclasts for active alveolar bone resorption. This is essential for the prompt timing of tooth eruption. The results also suggest the possibility that insufficient alveolar bone resorption is one of the cellular mechanisms of the delayed tooth eruption in CCD patients.

The methods used in these studies are different and may explain the disparity in the conclusions. Furthermore, it must be pointed out that in CCD the primary dentition is rarely affected. It may be assumed that the dentition in mice represents the human primary dentition, where the effects of haploinsufficiency on the eruption and formation of the human dentition may be too subtle to detect in the clinical situation. One major shortcoming of the murine model in this respect is that no secondary dentition develops.

Primary teeth erupt on time in CCD patients. Similarly, the permanent lower incisors and first molars generally erupt on time. However, the subsequent permanent teeth exhibit a delay in eruption, presumably as a result of defective eruption pathway formation (Fig. 1). There does not seem to be a close correlation between the number and positioning of supernumerary teeth and the delay in eruption (114, 115).

Similarly, birth length in CCD children is normal, but height drops below or around the second centile between 4 and 8 yr of age (116, 117). This pattern of development is reflected in the facial morphology, the characteristic frontal bossing and maxillary retrusion often not becoming evident until the later stages of childhood (118). Thus, the effect of haploinsufficiency on the craniofacial complex manifests late in a large number of cases. Whether this is caused by variation in RUNX2 isoform levels, or simply by a greater requirement for the gene product at that age, is unclear.

As the skeletal symptoms of CCD are not usually a social or physical handicap, and develop late, one of the factors leading to a diagnosis of CCD cases may be the observed anomalies in tooth eruption leading the patient to present for treatment (8).

Tooth eruption is controlled by precise osteoclast–osteoblast interaction. Osteoclastogenesis in the alveolar bone, which is essential for the accommodation of normal tooth development and eruption, is mediated by RANK–RANKL signaling (119). The spatiotemporal pattern and relative abundance of CSF-1, RANKL, and OPG during tooth eruption are key determinants of site-specific osteoclast activity in bone surrounding the tooth (120).

Runx2 is expressed in the alveolar bone at all stages of development and during tooth eruption (4,81). Evidence points to RUNX2 acting, either directly or indirectly, on the OPG/RANK/RANKL system to influence bone remodeling. Communication takes place between tooth germs and bone-forming/resorbing cells, synchronizing the two processes, perhaps to ensure correct spatial positioning of teeth in the jaws (82,121). The delayed and ectopic eruption of teeth seen in subjects with CCD may be caused by loss of function of RUNX2, both in respect of reduced CSF-1 production and disruption of the OPG/RANK/RANKL pathway.

Bone healing in CCD patients is not influenced. Normal healing has been reported after maxillofacial surgery (122), and osseointegration seems to be unaffected (123). Deciduous teeth are extracted, and bone is removed over unerupted teeth in order to encourage their eruption (124). It is possible that the inflammatory response to surgery may induce tooth eruption by re-activation of monocyte recruitment and osteoclast formation.

Supernumerary teeth

Supernumerary teeth are considered to be a diagnostic feature of CCD. However, the number of supernumerary teeth is variable, and several reports exist where no supernumeraries exist or hypodontia is reported (125–127).

Minor mutations in the highly conserved RHD are, in general, more likely to produce the classic CCD phenotype than the more robust flanking Q/A and PST domains. The phenotype in mutations involving the RHD is dependent on the residual transactivation potential of the protein (126, 128).

A dose-related effect seems to be present, as the milder cases of CCD, and those exhibiting primary dental anomalies, are associated with mutations that reduce, but do not abolish, protein stability, DNA binding, and transactivation. However, attempts to correlate the number of supernumerary teeth with the severity of skeletal symptoms are inconclusive. Furthermore, identical mutations produce different numbers of supernumerary teeth (126–130). All CCD mutations, including those which primarily feature dental anomalies, have highly variable phenotypic expression. This may indicate the overlying influence of other factors.

Runx2$^{-/-}$ and *Runx*2$^{+/-}$ mice were both found to exhibit lingual buds in front of the upper molars, and these were much more prominent than in wild-type mice (84,87). Furthermore, *Shh* signaling was not inhibited in the lingual buds of the knockout mice (87). Shh is necessary for tooth formation, both in the bud and cap stages (131), and its expression is spatially controlled to limit it to regions of tooth development (132).

These buds presumably represent the mouse secondary dentition, and it is likely that RUNX2 acts to inhibit formation of these buds. It may appear contradictory that the inhibition of RUNX2 function may arrest primary tooth development but stimulates the formation of secondary teeth. However, it is not unusual, during embryogenesis, for the same gene to have different effects at different developmental stages (87).

RUNX2 may normally function as a cell growth inhibitor in immature osteoblasts. It acts by supporting exit from the cell cycle, thus promoting increased expression of the osteoblast phenotype (133). Human *Runx2* $^{+/-}$ pulp cells proliferate at a far greater rate than their normal counterparts (98). This lends support to the theory that *Runx2* controls the proliferation of cells and may exert specific control on the dental lamina and the formation of successive dentitions.

It is easy to see how loss of function of this gene would encourage proliferation of the dental lamina. It is also easy to see how tooth eruption may be affected. However, the lack of correlation between the loss of RUNX2 function and the number of supernumerary teeth confuses matters. It is interesting to note that a mutation

affecting just the terminal VWRPY repressor motif produced a phenotype with only mildly hypoplastic clavicles and supernumerary teeth (5).

Concluding remarks

Runx2 is temporally and spatially regulated. The several signaling pathways that converge on this gene, and the existence of numerous splice variants with different N and C termini, substantiate its diverse actions on bone and tooth tissues.

One effect of a mutation may be to alter the proportions of the splice variants of the gene which will, in turn, affect its downstream pathways by altering the levels of interacting products. This would have different effects on different tissues. The inconclusive results of attempts to correlate the number of supernumerary teeth with eruption timing or skeletal effects would support this, as the action of RUNX2 on the dental lamina would be different to that on osteoblast function.

The dosage sensitivity exhibited helps to explain the large intrafamilial variability. Genetic or epigenetic modifiers may influence the phenotype, as may the transcriptional status of the unmutated allele (126).

There exist several cases of clinically diagnosed CCD where no mutation is detectable. Locus heterogeneity is a possible explanation for this; however, all CCD families tested map to the 6p21 locus (134–137).

Another explanation may be mutations within as-yet poorly characterized intronic or regulatory sequences. Variations in the promoter sequence point to a possible alternative mechanism for disruption of normal RUNX2 function (138, 139). Hypermethylation of the P2 promoter of the *Runx3* gene severely affects its function (140, 141). Given the high homology between these genes, the possibility of epigenetic effects on the *Runx2* regulatory regions should not be discounted.

References

1. Mundlos S, Otto F, Mundlos C, Mulliken JB, Aylsworth AS, Albright S, Lindhout D, Cole WG, Henn W, Knoll JH, Owen MJ, Mertelsmann R, Zabel BU, Olsen BR. Mutations involving the transcription factor CBFA1 cause cleidocranial dysplasia. *Cell* 1997; **89**: 773–779.
2. Komori T, Yagi H, Nomura S, Yamaguchi A, Sasaki K, Deguchi K, Shimizu Y, Bronson RT, Gao YH, Inada M, Sato M, Okamoto R, Kitamura Y, Yoshiki S, Kishimoto T. Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. *Cell* 1997; **89**: 755–764.
3. Ducy P, Zhang R, Geoffroy V, Ridall AL, Karsenty G. Osf2/Cbfa1: a transcriptional activator of osteoblast differentiation. *Cell* 1997; **89**: 747–754.
4. Ducy P, Starbuck M, Priemel M, Shen J, Pinero G, Geoffroy V, Amling M, Karsenty G. A Cbfa1-dependent genetic pathway controls bone formation beyond embryonic development. *Genes Dev* 1999; **13**: 1025–1036.
5. Quack I, Vonderstrass B, Stock M, Aylsworth AS, Becker A, Brueton L, Lee PJ, Majewski F, Mulliken JB, Suri M, Zenker M, Mundlos S, Otto F. Mutation analysis of core binding factor A1 in patients with cleidocranial dysplasia. *Am J Hum Genet* 1999; **65**: 1268–1278.
6. D'Souza RN, Aberg T, Gaikwad J, Cavender A, Owen M, Karsenty G, Thesleff I. Cbfa1 is required for epithelial–mesenchymal interactions regulating tooth development in mice. *Development* 1999; **126**: 2911–2920.
7. Hitchin AD. Cementum and other root abnormalities of permanent teeth in cleidocranial dysostosis. *Br Dent J* 1975; **139**: 313–318.
8. Jensen BL, Kreiborg S. Development of the dentition in cleidocranial dysplasia. *J Oral Pathol Med* 1990; **19**: 89–93.
9. Levanon D, Negreanu V, Bernstein Y, Bar-Am I, Avivi L, Groner Y. AML1, AML2, and AML3, the human members of the runt domain gene-family: cDNA structure, expression, and chromosomal localization. *Genomics* 1994; **23**: 425–432.
10. Xiao ZS, Thomas R, Hinson TK, Quarles LD. Genomic structure and isoform expression of the mouse, rat and human Cbfa1/Osf2 transcription factor. *Gene* 1998; **214**: 187–197.
11. Geoffroy V, Corral DA, Zhou L, Lee B, Karsenty G. Genomic organization, expression of the human CBFA1 gene, and evidence for an alternative splicing event affecting protein function. *Mamm Genome* 1998; **9**: 54–57.
12. Kundu M, Javed A, Jeon JP, Horner A, Shum L, Eckhaus M, Muenke M, Lian JB, Yang Y, Nuckolls GH, Stein GS, Liu PP. Cbfbeta interacts with Runx2 and has a critical role in bone development. *Nat Genet* 2002; **32**: 639–644.
13. Yoshida CA, Furuichi T, Fujita T, Fukuyama R, Kanatani N, Kobayashi S, Satake M, Takada K, Komori T. Core-binding factor beta interacts with Runx2 and is required for skeletal development. *Nat Genet* 2002; **32**: 633–638.
14. Ito Y. Molecular basis of tissue-specific gene expression mediated by the Runt domain transcription factor PEBP2/CBF. *Genes Cells* 1999; **4**: 685–696.
15. Huang G, Shigesada K, Ito K, Wee HJ, Yokomizo T, Ito Y. Dimerization with PEBP2beta protects RUNX1/AML1 from ubiquitin-proteasome-mediated degradation. *EMBO J* 2001; **20**: 723–733.
16. Young MF, Kerr JM, Ibaraki K, Heegaard AM, Robey PG. Structure, expression, and regulation of the major non-collagenous matrix proteins of bone. *Clin Orthop Relat Res* 1992; **281**: 275–294.
17. Robey PG, Fedarko NS, Hefferan TE, Bianco P, Vetter UK, Grzesik W, Friedenstein A, Van der PG, Mintz KP, Young MF. Structure and molecular regulation of bone matrix proteins. *J Bone Miner Res* 1993; **8**: S483–S487.
18. Mackie EJ. Osteoblasts: novel roles in orchestration of skeletal architecture. *Int J Biochem Cell Biol* 2003; **35**: 1301–1305.
19. Young MF. Bone matrix proteins: more than markers. *Calcif Tissue Int* 2003; **72**: 2–4.
20. Horowitz M. Matrix proteins versus cytokines in the regulation of osteoblast function and bone formation. *Calcif Tissue Int* 2003; **72**: 5–7.
21. Otto F, Thornell AP, Crompton T, Denzel A, Gilmour KC, Rosewell IR, Stamp GW, Beddington RS, Mundlos S, Olsen BR, Selby PB, Owen MJ. Cbfa1, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development. *Cell* 1997; **89**: 765–771.
22. Shevde NK, Bendixen AC, Maruyama M, Li BL, Billmire DA. Enhanced activity of osteoblast differentiation factor (PEBP2alphaA2/CBFa1) in affected sutural osteoblasts from patients with nonsyndromic craniosynostosis. *Cleft Palate Craniofac J* 2001; **38**: 606–614.
23. Geoffroy V, Kneissel M, Fournier B, Boyde A, Matthias P. High bone resorption in adult aging transgenic mice overexpressing Cbfa1/Runx2 in cells of the osteoblastic lineage. *Mol Cell Biol* 2002; **22**: 6222–6233.
24. Liu W, Toyosawa S, Furuichi T, Kanatani N, Yoshida C, Liu Y, Himeno M, Narai S, Yamaguchi A, Komori T. Overexpression of Cbfa1 in osteoblasts inhibits osteoblast maturation and causes osteopenia with multiple fractures. *J Cell Biol* 2001; **155**: 157–166.
25. Drissi H, Luc Q, Shakoori R, Chuva De Sousa LS, Choi JY, Terry A, Hu M, Jones S, Neil JC, Lian JB, Stein JL, van Wijnen AJ, Stein GS. Transcriptional autoregulation of

the bone related CBFA1/RUNX2 gene. *J Cell Physiol* 2000; **184**: 341–350.
26. Christiansen M, Kveiborg M, Kassem M, Clark BFC, Rattan SIS. CBFA1 and topoisomerase I mRNA levels decline during cellular aging of human trabecular osteoblasts. *J Gerontol A Biol Sci Med Sci* 2000; **55**: B194–B200.
27. Franceschi RT, Xiao G, Jiang D, Gopalakrishnan R, Yang S, Reith E. Multiple signaling pathways converge on the Cbfa1/Runx2 transcription factor to regulate osteoblast differentiation. *Connect Tissue Res* 2003; **44**: 109–116.
28. Franceschi RT, Xiao G. Regulation of the osteoblast-specific transcription factor, Runx2: responsiveness to multiple signal transduction pathways. *J Cell Biochem* 2003; **88**: 446–454.
29. Bae SC, Lee YH. Phosphorylation, acetylation and ubiquitination: the molecular basis of RUNX regulation. *Gene* 2006; **366**: 58–66.
30. Zhao M, Qiao M, Oyajobi BO, Mundy GR, Chen D. E3 ubiquitin ligase Smurf1 mediates core-binding factor {alpha}1/Runx2 degradation and plays a specific role in osteoblast differentiation. *J Biol Chem* 2003; **278**: 27939–27944.
31. Jeon EJ, Lee KY, Choi NS, Lee MH, Kim HN, Jin YH, Ryoo HM, Choi JY, Yoshida M, Nishino N, Oh BC, Lee KS, Lee YH, Bae SC. Bone morphogenetic protein-2 stimulates RUNX2 acetylation. *J Biol Chem* 2006; **281**; 16502–16511.
32. Shui C, Spelsberg TC, Riggs BL, Khosla S. Changes in Runx2/Cbfa1 expression and activity during osteoblastic differentiation of human bone marrow stromal cells. *J Bone Miner Res* 2003; **18**: 213–221.
33. Ducy P, Karsenty G. Two distinct osteoblast-specific cis-acting elements control expression of a mouse osteocalcin gene. *Mol Cell Biol* 1995; **15**: 1858–1869.
34. Ziros PG, Gil AP, Georgakopoulos T, Habeos I, Kletsas D, Basdra EK, Papavassiliou AG. The bone-specific transcriptional regulator Cbfa1 is a target of mechanical signals in osteoblastic cells. *J Biol Chem* 2002; **277**: 23934–23941.
35. Zaidi SK, Sullivan AJ, van Wijnen AJ, Stein JL, Stein GS, Lian JB. Integration of Runx and Smad regulatory signals at transcriptionally active subnuclear sites. *Proc Natl Acad Sci USA* 2002; **99**: 8048–8053.
36. Thirunavukkarasu K, Mahajan M, McLarren KW, Stifani S, Karsenty G. Two domains unique to osteoblast-specific transcription factor Osf2/Cbfa1 contribute to its transactivation function and its inability to heterodimerize with Cbfbeta. *Mol Cell Biol* 1998; **18**: 4197–4208.
37. Albrecht AN, Kornak U, Boddrich A, Suring K, Robinson PN, Stiege AC, Lurz R, Stricker S, Wanker EE, Mundlos S. A molecular pathogenesis for transcription factor associated poly-alanine tract expansions. *Hum Mol Genet* 2004; **13**: 2351–2359.
38. Zhang YW, Yasui N, Ito K, Huang G, Fujii M, Hanai J, Nogami H, Ochi T, Miyazono K, Ito Y. A RUNX2/PEBP2alpha A/CBFA1 mutation displaying impaired transactivation and Smad interaction in cleidocranial dysplasia. *Proc Natl Acad Sci USA* 2000; **97**: 10549–10554.
39. Sierra J, Villagra A, Paredes R, Cruzat F, Gutierrez S, Javed A, Arriagada G, Olate J, Imschenetzky M, van Wijnen AJ, Lian JB, Stein GS, Stein JL, Montecino M. Regulation of the bone-specific osteocalcin gene by p300 requires Runx2/Cbfa1 and the vitamin D3 receptor but not p300 intrinsic histone acetyltransferase activity. *Mol Cell Biol* 2003; **23**: 3339–3351.
40. Westendorf JJ, Zaidi SK, Cascino JE, Kahler R, van Wijnen AJ, Lian JB, Yoshida M, Stein GS, Li X. Runx2 (Cbfa1, AML-3) interacts with histone deacetylase 6 and represses the p21 (CIP1/WAF1) promoter. *Mol Cell Biol* 2002; **22**: 7982–7992.
41. McLarren KW, Theriault FM, Stifani S. Association with the nuclear matrix and interaction with Groucho and RUNX proteins regulate the transcription repression activity of the basic helix loop helix factor Hes1. *J Biol Chem* 2001; **276**: 1578–1584.
42. Javed A, Guo B, Hiebert S, Choi JY, Green J, Zhao SC, Osborne MA, Stifani S, Stein JL, Lian JB, van Wijnen AJ, Stein GS. Groucho/TLE/R-esp proteins associate with the nuclear matrix and repress RUNX (CBF (alpha)/AML/PEBP2 (alpha) dependent activation of tissue-specific gene transcription. *J Cell Sci* 2000; **113**: 2221–2231.
43. Fenrick R, Amann JM, Lutterbach B, Wang L, Westendorf JJ, Downing JR, Hiebert SW. Both TEL and AML-1 contribute repression domains to the t(12;21) fusion protein. *Mol Cell Biol* 1999; **19**: 6566–6574.
44. Zaidi SK, Sullivan AJ, Medina R, Ito Y, van Wijnen AJ, Stein JL, Lian JB, Stein GS. Tyrosine phosphorylation controls Runx2-mediated subnuclear targeting of YAP to repress transcription. *EMBO J* 2004; **23**: 790–799.
45. Westendorf JJ. Transcriptional co-repressors of Runx2. *J Cell Biochem* 2006; **98**: 54–64.
46. Wang W, Wang YG, Reginato AM, Glotzer DJ, Fukai N, Plotkina S, Karsenty G, Olsen BR. Groucho homologue Grg5 interacts with the transcription factor Runx2-Cbfa1 and modulates its activity during postnatal growth in mice. *Dev Biol* 2004; **270**: 364–381.
47. Schroeder TM, Kahler RA, Li X, Westendorf JJ. Histone deacetylase 3 interacts with runx2 to repress the osteocalcin promoter and regulate osteoblast differentiation. *J Biol Chem* 2004; **279**: 41998–42007.
48. Vega RB, Matsuda K, Oh J, Barbosa AC, Yang X, Meadows E, McAnally J, Pomajzl C, Shelton JM, Richardson JA, Karsenty G, Olson EN. Histone deacetylase 4 controls chondrocyte hypertrophy during skeletogenesis. *Cell* 2004; **119**: 555–566.
49. Pelletier N, Champagne N, Stifani S, Yang XJ. MOZ and MORF histone acetyltransferases interact with the Runt-domain transcription factor Runx2. *Oncogene* 2002; **21**: 2729–2740.
50. Thomas DM, Johnson SA, Sims NA, Trivett MK, Slavin JL, Rubin BP, Waring P, McArthur GA, Walkley CR, Holloway AJ, Diyagama D, Grim JE, Clurman BE, Bowtell DD, Lee JS, Gutierrez GM, Piscopo DM, Carty SA, Hinds PW. Terminal osteoblast differentiation, mediated by runx2 and p27KIP1, is disrupted in osteosarcoma. *J Cell Biol* 2004; **167**: 925–934.
51. Cui CB, Cooper LF, Yang X, Karsenty G, Aukhil I. Transcriptional coactivation of bone-specific transcription factor Cbfa1 by TAZ. *Mol Cell Biol* 2003; **23**: 1004–1013.
52. Bialek P, Kern B, Yang X, Schrock M, Sosic D, Hong N, Wu H, Yu K, Ornitz DM, Olson EN, Justice MJ, Karsenty G. A twist code determines the onset of osteoblast differentiation. *Dev Cell* 2004; **6**: 423–435.
53. Harrington KS, Javed A, Drissi H, McNeil S, Lian JB, Stein JL, van Wijnen AJ, Wang YL, Stein GS. Transcription factors RUNX1/AML1 and RUNX2/Cbfa1 dynamically associate with stationary subnuclear domains. *J Cell Sci* 2002; **115**: 4167–4176.
54. Zeng C, McNeil S, Pockwinse S, Nickerson J, Shopland L, Lawrence JB, Penman S, Hiebert S, Lian JB, van Wijnen AJ, Stein JL, Stein GS. Intranuclear targeting of AML/CBFalpha regulatory factors to nuclear matrix-associated transcriptional domains. *Proc Natl Acad Sci USA* 1998; **95**: 1585–1589.
55. Zaidi SK, Javed A, Choi JY, van Wijnen AJ, Stein JL, Lian JB, Stein GS. A specific targeting signal directs Runx2/Cbfa1 to subnuclear domains and contributes to transactivation of the osteocalcin gene. *J Cell Sci* 2001; **114**: 3093–3102.
56. Javed A, Barnes GL, Pratap J, Antkowiak T, Gerstenfeld LC, van Wijnen AJ, Stein JL, Lian JB, Stein GS. Impaired intranuclear trafficking of Runx2 (AML3/CBFA1) transcription factors in breast cancer cells inhibits osteolysis in vivo. *Proc Natl Acad Sci USA* 2005; **102**: 1454–1459.
57. Xiao ZS, Hinson TK, Quarles LD. Cbfa1 isoform overexpression upregulates osteocalcin gene expression in non-osteoblastic and pre-osteoblastic cells. *J Cell Biochem* 1999; **74**: 596–605.
58. Bae SC, Ogawa E, Maruyama M, Oka H, Satake M, Shigesada K, Jenkins NA, Gilbert DJ, Copeland NG, Ito Y. PEBP2 alpha B/mouse AML1 consists of multiple isoforms

that possess differential transactivation potentials. *Mol Cell Biol* 1994; **14**: 3242–3252.
59. Banerjee C, Javed A, Choi JY, Green J, Rosen V, van Wijnen AJ, Stein JL, Lian JB, Stein GS. Differential regulation of the two principal Runx2/Cbfa1 N-terminal isoforms in response to bone morphogenetic protein-2 during development of the osteoblast phenotype. *Endocrinology* 2001; **142**: 4026–4039.
60. Harada H, Tagashira S, Fujiwara M, Ogawa S, Katsumata T, Yamaguchi A, Komori T, Nakatsuka M. Cbfa1 isoforms exert functional differences in osteoblast differentiation. *J Biol Chem* 1999; **274**: 6972–6978.
61. Park MH, Shin HI, Choi JY, Nam SH, Kim YJ, Kim HJ, Ryoo HM. Differential expression patterns of Runx2 isoforms in cranial suture morphogenesis. *J Bone Miner Res* 2001; **16**: 885–892.
62. Xiao ZS, Hjelmeland AB, Quarles LD. Selective deficiency of the 'bone-related' Runx2-II unexpectedly preserves osteoblast-mediated skeletogenesis. *J Biol Chem* 2004; **279**: 20307–20313.
63. Choi KY, Lee SW, Park MH, Bae YC, Shin HI, Nam S, Kim YJ, Kim HJ, Ryoo HM. Spatio-temporal expression patterns of Runx2 isoforms in early skeletogenesis. *Exp Mol Med* 2002; **34**: 426–433.
64. Choi KY, Kim HJ, Lee MH, Kwon TG, Nah HD, Furuichi T, Komori T, Nam SH, Kim YJ, Kim HJ, Ryoo HM. Runx2 regulates FGF2-induced Bmp2 expression during cranial bone development. *Dev Dyn* 2005; **233**: 115–121.
65. Lee MH, Kim YJ, Yoon WJ, Kim JI, Kim BG, Hwang YS, Wozney JM, Chi XZ, Bae SC, Choi KY, Cho JY, Choi JY, Ryoo HM. Dlx5 specifically regulates Runx2 type II expression by binding to homeodomain-response elements in the Runx2 distal promoter. *J Biol Chem* 2005; **280**: 35579–35587.
66. Chen S, Gu TT, Sreenath T, Kulkarni AB, Karsenty G, MacDougall M. Spatial expression of Cbfa1/Runx2 isoforms in teeth and characterization of binding sites in the DSPP gene. *Connect Tissue Res* 2002; **43**: 338–344.
67. Kobayashi I, Kiyoshima T, Wada H, Matsuo K, Nonaka K, Honda JY, Koyano K, Sakai H. Type II/III Runx2/Cbfa1 is required for tooth germ development. *Bone* 2006; **38**: 836–844.
68. Xiao ZS, Simpson LG, Quarles LD. IRES-dependent translational control of Cbfa1/Runx2 expression. *J Cell Biochem* 2003; **88**: 493–505.
69. Stewart M, Terry A, Hu M, O'Hara M, Blyth K, Baxter E, Cameron E, Onions DE, Neil JC. Proviral insertions induce the expression of bone-specific isoforms of PEBP2alphaA (CBFA1): evidence for a new myc collaborating oncogene. *Proc Natl Acad Sci USA* 1997; **94**: 8646–8651.
70. Terry A, Kilbey A, Vaillant F, Stewart M, Jenkins A, Cameron E, Neil JC. Conservation and expression of an alternative 3′ exon of Runx2 encoding a novel proline-rich C-terminal domain. *Gene* 2004; **336**: 115–125.
71. Manolagas SC. Birth and death of bone cells: basic regulatory mechanisms and implications for the pathogenesis and treatment of osteoporosis. *Endocr Rev* 2000; **21**: 115–137.
72. Gori F, Hofbauer LC, Dunstan CR, Spelsberg TC, Khosla S, Riggs BL. The expression of osteoprotegerin and RANK ligand and the support of osteoclast formation by stromal-osteoblast lineage cells is developmentally regulated. *Endocrinology* 2000; **141**: 4768–4776.
73. Gao YH, Shinki T, Yuasa T, Kataoka-Enomoto H, Komori T, Suda T, Yamaguchi A. Potential role of cbfa1, an essential transcriptional factor for osteoblast differentiation, in osteoclastogenesis: regulation of mRNA expression of osteoclast differentiation factor (ODF). *Biochem Biophys Res Commun* 1998; **252**: 697–702.
74. Enomoto H, Shiojiri S, Hoshi K, Furuichi T, Fukuyama R, Yoshida CA, Kanatani N, Nakamura R, Mizuno A, Zanma A, Yano K, Yasuda H, Higashio K, Takada K, Komori T. Induction of osteoclast differentiation by Runx2 through receptor activator of nuclear factor-kappa B ligand (RANKL) and osteoprotegerin regulation and partial rescue of osteoclastogenesis in Runx2-/- mice by RANKL transgene. *J Biol Chem* 2003; **278**: 23971–23977.
75. Thirunavukkarasu K, Halladay DL, Miles RR, Yang X, Galvin RJ, Chandrasekhar S, Martin TJ, Onyia JE. The osteoblast-specific transcription factor Cbfa1 contributes to the expression of osteoprotegerin, a potent inhibitor of osteoclast differentiation and function. *J Biol Chem* 2000; **275**: 25163–25172.
76. Thirunavukkarasu K, Miles RR, Halladay DL, Yang X, Galvin RJS, Chandrasekhar S, Martin TJ, Onyia JE. Stimulation of osteoprotegerin (OPG) gene expression by transforming growth factor-beta (TGF-beta). Mapping of the OPG promoter region that mediates TGF-beta effects. *J Biol Chem* 2001; **276**: 36241–36250.
77. Kitazawa R, Kitazawa S, Maeda S. Promoter structure of mouse RANKL/TRANCE/OPGL/ODF gene. *Biochim Biophys Acta* 1999; **1445**: 134–141.
78. O'Brien CA, Kern B, Gubrij I, Karsenty G, Manolagas SC. Cbfa1 does not regulate RANKL gene activity in stromal/osteoblastic cells. *Bone* 2002; **30**: 453–462.
79. Notoya M, Otsuka E, Yamaguchi A, Hagiwara H. Runx-2 is not essential for the vitamin D-regulated expression of RANKL and osteoprotegerin in osteoblastic cells. *Biochem Biophys Res Commun* 2004; **324**: 655–660.
80. Blair HC, Robinson LJ, Zaidi M. Osteoclast signalling pathways. *Biochem Biophys Res Commun* 2005; **328**: 728–738.
81. Jiang H, Sodek J, Karsenty G, Thomas H, Ranly D, Chen J. Expression of core binding factor Osf2/Cbfa-1 and bone sialoprotein in tooth development. *Mech Dev* 1999; **81**: 169–173.
82. Ohazama A, Courtney JM, Sharpe PT. Opg, Rank, and Rankl in tooth development: co-ordination of odontogenesis and osteogenesis. *J Dent Res* 2004; **83**: 241–244.
83. Aberg T, Wang XP, Kim JH, Yamashiro T, Bei M, Rice R, Ryoo HM, Thesleff I. Runx2 mediates FGF signaling from epithelium to mesenchyme during tooth morphogenesis. *Dev Biol* 2004; **270**: 76–93.
84. Aberg T, Cavender A, Gaikwad JS, Bronckers ALJJ, Wang X, Waltimo-Siren J, Thesleff I, D'Souza RN. Phenotypic changes in dentition of Runx2 homozygote-null mutant mice. *J Histochem Cytochem* 2004; **52**: 131–140.
85. Thomas BL, Tucker AS, Qui M, Ferguson CA, Hardcastle Z, Rubenstein JL, Sharpe PT. Role of Dlx-1 and Dlx-2 genes in patterning of the murine dentition. *Development* 1997; **124**: 4811–4818.
86. Ferguson CA, Tucker AS, Christensen L, Lau AL, Matzuk MM, Sharpe PT. Activin is an essential early mesenchymal signal in tooth development that is required for patterning of the murine dentition. *Genes Dev* 1998; **12**: 2636–2649.
87. Wang XP, Aberg T, James MJ, Levanon D, Groner Y, Thesleff I. Runx2 (Cbfa1) inhibits Shh signaling in the lower but not upper molars of mouse embryos and prevents the budding of putative successional teeth. *J Dent Res* 2005; **84**: 138–143.
88. Bronckers AL, Engelse MA, Cavender A, Gaikwad J, D'Souza RN. Cell-specific patterns of Cbfa1 mRNA and protein expression in postnatal murine dental tissues. *Mech Dev* 2001; **101**: 255–258.
89. Dhamija S, Krebsbach PH. Role of Cbfa1 in ameloblastin gene transcription. *J Biol Chem* 2001; **276**: 35159–35164.
90. Fukuta Y, Totsuka M, Fukuta Y, Takeda Y, Yoshida Y, Niitsu J, Yamamoto H. Histological and analytical studies of a tooth in a patient with cleidocranial dysostosis. *J Oral Sci* 2001; **43**: 85–89.
91. Yamamoto H, Sakae T, Davies JE. Cleidocranial dysplasia: a light microscope, electron microscope, and crystallographic study. *Oral Surg Oral Med Oral Pathol* 1989; **68**: 195–200.
92. Lukinmaa PL, Jensen BL, Thesleff I, Andreasen JO, Kreiborg S. Histological observations of teeth and peridental tissues in cleidocranial dysplasia imply increased activity of odontogenic epithelium and abnormal bone remodeling. *J Craniofac Genet Dev Biol* 1995; **15**: 212–221.

93. Qin C, Brunn JC, Cadena E, Ridall A, Tsujigiwa H, Nagatsuka H, Nagai N, Butler WT. The expression of dentin sialophosphoprotein gene in bone. *J Dent Res* 2002; **81**: 392–394.
94. Baba O, Qin C, Brunn JC, Jones JE, Wygant JN, McIntyre BW, Butler WT. Detection of dentin sialoprotein in rat periodontium. *Eur J Oral Sci* 2004; **112**: 163–170.
95. Chen S, Rani S, Wu Y, Unterbrink A, Gu TT, Gluhak-Heinrich J, Chuang HH, MacDougall M. Differential regulation of dentin sialophosphoprotein expression by Runx2 during odontoblast cytodifferentiation. *J Biol Chem* 2005; **280**: 29717–29727.
96. Narayanan K, Srinivas R, Ramachandran A, Hao J, Quinn B, George A. Differentiation of embryonic mesenchymal cells to odontoblast-like cells by overexpression of dentin matrix protein 1. *Proc Natl Acad Sci USA* 2001; **98**: 4516–4521.
97. Fen JQ, Zhang J, Dallas SL, Lu Y, Chen S, Tan X, Owen M, Harris SE, MacDougall M. Dentin matrix protein 1, a target molecule for Cbfa1 in bone, is a unique bone marker gene. *J Bone Miner Res* 2002; **17**: 1822–1831.
98. Chen S, Santos L, Wu Y, Vuong R, Gay I, Schulze J, Chuang HH, MacDougall M. Altered gene expression in human cleidocranial dysplasia dental pulp cells. *Arch Oral Biol* 2005; **50**: 227–236.
99. Liu H, Li W, Shi S, Habelitz S, Gao C, Denbesten P. MEPE is downregulated as dental pulp stem cells differentiate. *Arch Oral Biol* 2005; **50**: 923–928.
100. Yamashiro T, Aberg T, Levanon D, Groner Y, Thesleff I. Expression of Runx1, -2 and -3 during tooth, palate and craniofacial bone development. *Mech Dev* 2002; **119**: S107–S110.
101. Gaikwad JS, Hoffmann M, Cavender A, Bronckers AL, D'Souza RN. Molecular insights into the lineage-specific determination of odontoblasts: the role of Cbfa1. *Adv Dent Res* 2001; **15**: 19–24.
102. Bosshardt DD. Are cementoblasts a subpopulation of osteoblasts or a unique phenotype? *J Dent Res* 2005; **84**: 390–406.
103. Dard M. Histology of alveolar bone and primary tooth roots in a case of cleidocranial dysplasia. *Bull Group Int Rech Sci Stomatol Odontol* 1993; **36**: 101–107.
104. Zou SJ, D'Souza RN, Ahlberg T, Bronckers AL. Tooth eruption and cementum formation in the Runx2/Cbfa1 heterozygous mouse. *Arch Oral Biol* 2003; **48**: 673–677.
105. Chung CR, Tsuji K, Nifuji A, Komori T, Soma K, Noda M. Micro-CT evaluation of tooth, calvaria and mechanical stress-induced tooth movement in adult Runx2/Cbfa1 heterozygous knock-out mice. *J Med Dent Sci* 2004; **51**: 105–113.
106. Seow WK, Hertzberg J. Dental development and molar root length in children with cleidocranial dysplasia. *Pediatr Dent* 1995; **17**: 101–105.
107. Hitchin AD, Fairley JM. Dental management in cleidocranial dysostosis. *Br J Oral Surg* 1974; **12**: 46–55.
108. McNamara CM, O'Riordan BC, Blake M, Sandy JR. Cleidocranial dysplasia: radiological appearances on dental panoramic radiography. *Dentomaxillofac Radiol* 1999; **28**: 89–97.
109. Saito M, Iwase M, Maslan S, Nozaki N, Yamauchi M, Handa K, Takahashi O, Sato S, Kawase T, Teranaka T, Narayanan AS. Expression of cementum-derived attachment protein in bovine tooth germ during cementogenesis. *Bone* 2001; **29**: 242–248.
110. Kato C, Kojima T, Komaki M, Mimori K, Duarte WR, Takenaga K, Ishikawa I. S100A4 inhibition by RNAi up-regulates osteoblast related genes in periodontal ligament cells. *Biochem Biophys Res Commun* 2004; **326**: 147–153.
111. Yang YQ, Li XT, Rabie AB, Fu MK, Zhang D. Human periodontal ligament cells express osteoblastic phenotypes under intermittent force loading in vitro. *Front Biosci* 2006; **11**: 776–781.
112. Shaikh R, Shusterman S. Delayed dental maturation in cleidocranial dysplasia. *ASDC J Dent Child* 1998; **65**: 355.
113. Yoda S, Suda N, Kitahara Y, Komori T, Ohyama K. Delayed tooth eruption and suppressed osteoclast number in the eruption pathway of heterozygous Runx2/Cbfa1 knockout mice. *Arch Oral Biol* 2004; **49**: 435–442.
114. Richardson A, Deussen FF. Facial and dental anomalies in cleidocranial dysplasia: a study of 17 cases. *Int J Paediatr Dent* 1994; **4**: 225–231.
115. Jensen BL, Kreiborg S. Dental treatment strategies in cleidocranial dysplasia. *Br Dent J* 1992; **172**: 243–247.
116. Jensen BL. Somatic development in cleidocranial dysplasia. *Am J Med Genet* 1990; **35**: 69–74.
117. Levin EJ, Sonnenschein H. Cleidocranial dysostosis. *N Y State J Med* 1963; **63**: 1562–1566.
118. Ishii K, Nielsen IL, Vargervik K. Characteristics of jaw growth in cleidocranial dysplasia. *Cleft Palate Craniofac J* 1998; **35**: 161–166.
119. Suzuki T, Suda N, Ohyama K. Osteoclastogenesis during mouse tooth germ development is mediated by receptor activator of NFKappa-B ligand (RANKL). *J Bone Miner Metab* 2004; **22**: 185–191.
120. Heinrich J, Bsoul S, Barnes J, Woodruff K, Abboud S. CSF-1, RANKL and OPG regulate osteoclastogenesis during murine tooth eruption. *Arch Oral Biol* 2005; **50**: 897–908.
121. Gao Y, Kobayashi H, Ganss B. The human KROX-26/ZNF22 gene is expressed at sites of tooth formation and maps to the locus for permanent tooth agenesis (He-Zhao deficiency). *J Dent Res* 2003; **82**: 1002–1007.
122. Dann JJ III, Crump P, Ringenberg QM. Vertical maxillary deficiency with cleidocranial dysplasia. Diagnostic findings and surgical-orthodontic correction. *Am J Orthod* 1980; **78**: 564–574.
123. Petropoulos VC, Balshi TJ, Balshi SF, Wolfinger GJ. Treatment of a patient with cleidocranial dysplasia using osseointegrated implants: a patient report. *Int J Oral Maxillofac Implants* 2004; **19**: 282–287.
124. Becker A, Lustmann J, Shteyer A. Cleidocranial dysplasia: Part 1 – General principles of the orthodontic and surgical treatment modality. *Am J Orthod Dentofacial Orthop* 1997; **111**: 28–33.
125. Hardt N. Osteo-dental dysplasia with poly- and hypodontia in cleidocranial dysostosis. *Quintessenz* 1973; **24**: 105–111.
126. Baumert U, Golan I, Redlich M, Aknin JJ, Muessig D. Cleidocranial dysplasia: molecular genetic analysis and phenotypic-based description of a Middle European patient group. *Am J Med Genet* 2005; **139**: 78–85.
127. Yoshida T, Kanegane H, Osato M, Yanagida M, Miyawaki T, Ito Y, Shigesada K. Functional analysis of RUNX2 mutations in Japanese patients with cleidocranial dysplasia demonstrates novel genotype-phenotype correlations. *Am J Hum Genet* 2002; **71**: 724–738.
128. Yoshida T, Kanegane H, Osato M, Yanagida M, Miyawaki T, Ito Y, Shigesada K. Functional analysis of RUNX2 mutations in cleidocranial dysplasia: novel insights into genotype-phenotype correlations. *Blood Cells Mol Dis* 2003; **30**: 184–193.
129. Tessa A, Salvi S, Casali C, Garavelli L, Digilio MC, Dotti MT, Di Giandomenico S, Valoppi M, Grieco GS, Comanducci G, Bianchini G, Fortini D, Federico A, Giannotti A, Santorelli FM. Six novel mutations of the RUNX2 gene in Italian patients with cleidocranial dysplasia. *Hum Mutat* 2003; **22**: 104.
130. Zhou G, Chen Y, Zhou L, Thirunavukkarasu K, Hecht J, Chitayat D, Gelb BD, Pirinen S, Berry SA, Greenberg CR, Karsenty G, Lee B. CBFA1 mutation analysis and functional correlation with phenotypic variability in cleidocranial dysplasia. *Hum Mol Genet* 1999; **8**: 2311–2316.
131. Cobourne MT, Hardcastle Z, Sharpe PT. Sonic hedgehog regulates epithelial proliferation and cell survival in the developing tooth germ. *J Dent Res* 2001; **80**: 1974–1979.
132. Cobourne MT, Miletich I, Sharpe PT. Restriction of sonic hedgehog signalling during early tooth development. *Development* 2004; **131**: 2875–2885.
133. Pratap J, Galindo M, Zaidi SK, Vradii D, Bhat BM, Robinson JA, Choi JY, Komori T, Stein JL, Lian JB, Stein GS, van Wijnen AJ. Cell growth regulatory role of Runx2

during proliferative expansion of preosteoblasts. *Cancer Res* 2003; **63**: 5357–5362.

134. Ramesar RS, Greenberg J, Martin R, Goliath R, Bardien S, Mundlos S, Beighton P. Mapping of the gene for cleidocranial dysplasia in the historical Cape Town (Arnold) kindred and evidence for locus homogeneity. *J Med Genet* 1996; **33**: 511–514.
135. Gelb BD, Cooper E, Shevell M, Desnick RJ. Genetic mapping of the cleidocranial dysplasia (CCD) locus on chromosome band 6p21 to include a microdeletion. *Am J Med Genet* 1995; **58**: 200–205.
136. Feldman GJ, Robin NH, Brueton LA, Robertson E, Thompson EM, Siegel-Bartelt J, Gasser DL, Bailey LC, Zackai EH, Muenke M. A gene for cleidocranial dysplasia maps to the short arm of chromosome 6. *Am J Hum Genet* 1995; **56**: 938–943.
137. Mundlos S, Mulliken JB, Abramson DL, Warman ML, Knoll JH, Olsen BR. Genetic mapping of cleidocranial dysplasia and evidence of a microdeletion in one family. *Hum Mol Genet* 1995; **4**: 71–75.
138. Napierala D, Garcia-Rojas X, Sam K, Wakui K, Chen C, Mendoza-Londono R, Zhou G, Zheng Q, Lee B. Mutations and promoter SNPs in RUNX2, a transcriptional regulator of bone formation. *Mol Genet Metab* 2005; **86**: 257–268.
139. Doecke JD, Day CJ, Stephens AS, Carter SL, van Daal A, Kotowicz MA, Nicholson GC, Morrison NA. Association of functionally different RUNX2 P2 promoter alleles with BMD. *J Bone Miner Res* 2006; **21**: 265–273.
140. Bae SC, Choi JK. Tumor suppressor activity of RUNX3. *Oncogene* 2004; **23**: 4336–4340.
141. Li QL, Kim HR, Kim WJ, Choi JK, Lee YH, Kim HM, Li LS, Kim H, Chang J, Ito Y, Youl LK, Bae SC. Transcriptional silencing of the RUNX3 gene by CpG hypermethylation is associated with lung cancer. *Biochem Biophys Res Commun* 2004; **314**: 223–228.

Ectopic Maxillary Canines: Segregation Analysis and a Twin Study

S. Camilleri[1*,2], C.M. Lewis[3], and F. McDonald[1]

[1]Department of Orthodontics and [3]Division of Genetics and Molecular Medicine, Kings College London, Dental Institute, Floor 22, Guy's Tower, St Thomas St., London SE1 9RT, UK, and [2]Faculty of Dental Surgery, University of Malta, Malta, GC; *corresponding author, simon.camilleri@kcl.ac.uk, simon.camilleri@um.edu.mt

J Dent Res 87(6):580-583, 2008

ABSTRACT

The etiology of ectopic canines is controversial, with opinion divided as to a genetic or environmental mechanism. This study addressed the hypothesis that genetic factors play a role in the etiology of ectopic maxillary canines. Sixty-three probands were identified, and information on the dental status of 395 relatives was determined. Pedigrees were constructed and the Relative Risk calculated. Complex segregation analysis was carried out by means of the Pedigree Analysis Package. The best mathematical model obtained was a single dominant gene with autosomal transmission, incomplete penetrance, and highly variable expression. Only two of seven pairs of monozygotic twins were concordant for ectopic canines. This is consistent with environmental or epigenetic variables affecting the phenotype. The low concordance rate is consistent with the low penetrance determined by the segregation analysis and further supports the existence of environmental factors.

KEY WORDS: tooth eruption, ectopic, cuspid, segregation analysis, Maltese.

Received March 28, 2007; Last revision January 29, 2008; Accepted February 8, 2008

INTRODUCTION

The ectopic canine is the second most frequently impacted tooth after the third molar, appearing in 3% of the Western population (Ericson and Kurol, 1986).

The etiology is controversial. A genetic basis has been suggested (Bjerklin *et al.*, 1992; Peck *et al.*, 1994). Several studies have shown the etiology of ectopic canines to be genetic and associated with other genetically interrelated dental anomalies (Svinhufvud *et al.*, 1988; Bjerklin *et al.*, 1992; Pirinen *et al.*, 1996; Baccetti, 1998). The sex ratio shows a bias toward females (Becker *et al.*, 1981; Ericson and Kurol, 1988), similar to other dental anomalies of genetic origin (Rose, 1966; Davis, 1987). The racial variation, female preponderance, familial occurrence, and association with other dental anomalies imply a polygenic etiology (Kotsomitis and Freer, 1997).

Environmental factors have also been identified. Palatal displacement of the canine may be due to local environmental factors, such as anatomically anomalous or late-developing lateral incisor roots (Becker *et al.*, 1981; Chaushu *et al.*, 2002, 2003). Excess space in the dental arch has been implicated (Paschos *et al.*, 2005). There is evidence that both genetic and environmental factors may be involved. (Ely *et al.*, 2006). Alteration of the local environment by extraction of the deciduous canines will ameliorate the condition (Ericson and Kurol, 1988; Power and Short, 1993).

The prevalence of ectopic canines in Maltese schoolchildren is 4-5.5% (Camilleri, 1995; Camilleri *et al.*, unpublished observations). This is higher than previously published figures (Ericson and Kurol, 1986), possibly due to the 'founder effect', since the Maltese population has grown dramatically, from 17,000 in 1535 to over 300,000 today (Blouet, 2004). This relative abundance and familial clustering of cases prompted a study to test the hypothesis that ectopic canines have a genetic etiology.

METHODS

Ethical approval was obtained from the University of Malta Medical School Ethics Committee. Informed consent was obtained from all participants prior to inclusion in the study. Thirty-seven consecutive probands with ectopic maxillary canines were identified during routine clinical examination at the private practice of SC and at the School Dental Clinic, Floriana. The inclusion criteria were Maltese Citizenship and a positive history of ectopic (buccally or palatally displaced) canines or failure of eruption of the canine tooth by the age of 16. Individuals affected by a genetic syndrome likely to have an adverse influence on tooth eruption were excluded. A further set of 26 consecutive families was selected on the strength of two probands, to identify families with a stronger predisposition. Families with an affected twin were investigated even if the other twin was unaffected. Ectopic or missing teeth were identified by clinical examination, from existing records, or family history, and radiographic investigation was undertaken only where clinically necessary.

We analyzed the pedigree data to assess the familial risk of ectopic canines and

other related phenotypes, using the computed familial relative risk.

We used the segregation analysis program Pedigree Analysis Program (PAP) (Hasstedt, 2005) to model the inheritance of ectopic canines throughout the 63 pedigrees. Pedigree members were defined as affected or unaffected with ectopic canines (with no further clinical information used). Genetic models assumed a single locus model with penetrances and mutation frequency maximized for each model. Nested models were compared by the likelihood ratio test, where the difference in -2 ln (likelihood) has a chi-square distribution, with the degrees of freedom equal to the difference in the number of additional parameters fitted in the most general model. Sporadic (no genetic effect), recessive, and dominant models were compared with a co-dominant model. A polygenic model was also tested and compared with the dominant model. An ascertainment correction for the probands in each family was applied.

Five sets of monozygotic and six sets of dizygotic twins were included in the sample. A further two sets of monozygotic twins and one set of triplets were subsequently referred by a colleague. These were not included in the pedigree or segregation analyses; however, they were used in assessments of pairwise concordance.

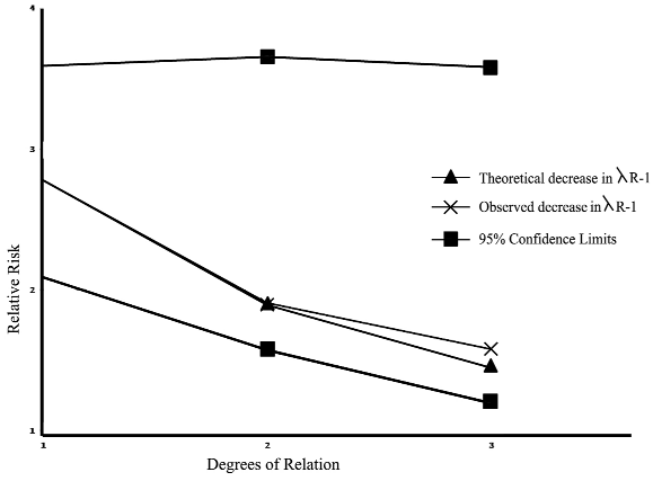


Figure 1. Relative risk drop-off. The reduction in Relative Risk against Degree of Relation is very similar to the theoretical curve for inheritance of a genetic disease. (λ_R = Relative Risk).

RESULTS

The number of individuals of known dental status was 524. One hundred and thirty-nine individuals had ectopic canines.

The percentage of dental anomalies in the sample for first-degree relatives was noted, and differences against published population prevalences were investigated. The prevalence of ectopic canines was significantly higher in first-degree relatives (15%, $p < 0.001$) compared with 4.4-5.5% for the Maltese population (Camilleri, 1995; Camilleri *et al.*, unpublished observations). Lateral incisor agenesis was also higher (7.88%, $p = 0.01$, as opposed to 3.21% for the general population) (Camilleri and Mulligan, 2007).

There were eight cases of maxillary canine transposition in the whole sample. Three were probands, five were first-degree relatives; one was a second-degree relative. Seven were in the upper jaw, giving a prevalence of 1.4%. This is significantly higher ($p < 0.001$) than the prevalence of 0.27% estimated in a Caucasian population (Yilmaz *et al.*, 2005). Two cases of mandibular canine-lateral incisor transposition were also recorded.

The percentages of first-, second-, and third-degree relatives with ectopic maxillary canines were used to calculate the relative risk (λ_R) to the first-, second-, and third-degree relatives. This is calculated as ($\lambda_R = \kappa_R/\kappa$) where κ is the population prevalence and κ_R is the percentage of relatives affected according to the type of relative, *i.e.,* of the first, second, or third degree (Table 1). Ascertainment was corrected for by exclusion of the probands from the calculation. We then used the relative risk to plot the drop-off for each decreasing degree of unilineal relationship (Fig. 1), which was close to the theoretical decrease for a genetic disease.

There was no difference in ectopic canine risk between families ascertained from one or two probands for first-, second-, or third-degree relatives ($p = 0.49$, $p = 0.52$, $p = 0.65$). Nor was there a difference in the numbers of sib-sib and parent-offspring affected pairs for each type of family ascertainment ($p = 0.89$).

There was an appreciable sex bias, with the affected Female-to-Male Ratio for first-, second-, and third-degree relatives being 1.78. The sex ratio for the probands was 1.64, and elimination of the probands from the calculation gave a ratio of 1.89. However, there was no appreciable risk difference between relatives of male or female probands ($p = 0.77$). The proportion of male first-degree relatives affected *vs*. the proportion of female first-degree relatives was not significant ($p = 0.12$).

Although there was no evidence of sex-linked transmission, 85% of the three-generation families showed instances where an obligate carrier showed a normal phenotype, although the condition was transmitted to their children (Fig. 2). There was no apparent pattern of augmentation or attenuation of phenotype through the generations.

We used the segregation analysis program PAP to fit different genetic models to the pedigrees, estimating parameters such as penetrance and mutation frequency. We applied a likelihood ratio test to determine any significant difference in the fit of the various models (Table 2). All the genetic models fitted provided a significantly better fit than the sporadic model, confirming that ectopic canines have a genetic basis. The -2log likelihood of the co-dominant model was very similar to the dominant model, but had more parameters. Therefore, the dominant model provided the more parsimonious model, with

Table 1. Relative Risks: Proportions of Ectopic Canine-affected Individuals Used for Calculation of the Relative Risks

	Relatives Affected with Ectopic Canines		Total Number of Individuals	Percentage of Relatives Affected	Relative Risk
	2 probands	1 proband			
1st-degree relatives	16	15	203	15.27	2.78
2nd-degree relatives	6	8	117	11.97	2.18
3rd-degree relatives	5	3	85	9.41	1.71
Average sampled family size	10	7			

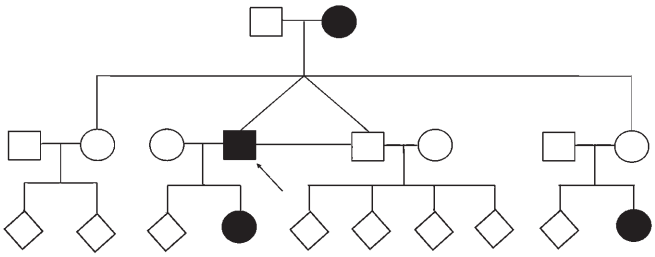


Figure 2. Ectopic canine pedigree. Pedigree showing autosomal-dominant transmission, incomplete penetrance, and discordant identical twins, with the trait transmitted to the offspring of the affected twin. The arrow indicates the proband.

the -2log (likelihood) for the recessive model being substantially higher. A mixed-dominant model did not provide a sufficiently improved fit, implying that there is no evidence for a polygenic component in addition to the Mendelian locus. These results show that a single dominant gene best accounts for the inheritance of ectopic canines in these pedigrees, with a mutation frequency of 11% and penetrances of 36% in mutation carriers.

From the dominant model fitted in the segregation analysis, the risk to the sibling of an ectopic canine case was calculated to be 0.18. This is similar to the estimate of 0.15 obtained from the pedigree analysis (RR 2.78 x population prevalence 5.5%). The population prevalence predicted by the dominant model was determined according to the formula: $q^2 * p_1 + 2\ q\ (1\text{-}q)\ p_2 + (1\text{-}q)^2\ p_3$, where q is the allele frequency, and p_1, p_2, and p_3 are the penetrances for the three genotypes. The calculated prevalence was 7%, which also compared well with the estimated population prevalence of 5.5%. This supports the dominant model as being the most likely.

Seven pairs of monozygotic twins were found. Of these, five were completely discordant, in that one twin was unaffected (Fig. 2). Of the other two pairs, one pair was mirror-image, and the other set had one twin affected unilaterally, the other bilaterally. The pairwise concordance in this sample was 28.6%. Seven pairs of dizygotic twins are also on record, with two showing concordance for ectopic canines, a pairwise concordance rate of 28.6%.

DISCUSSION

Dental Anomalies

There was a significantly higher prevalence of ectopic canines in first-degree relatives of ectopic canine probands (15%, $p < 0.001$) as compared with figures for the Maltese population (Camilleri, 1995; Camilleri *et al.*, unpublished observations). Familial clustering is a common feature of the trait; however, this could be due to environmental as well as genetic factors.

The prevalence of canine transposition in the sample was far higher than the 0.27% prevalence estimated in a Caucasian population (Yilmaz *et al.*, 2005). The highly significant difference is indicative of transposition, being an extreme variant of ectopic canine, and is consistent with the highly variable expression exhibited.

The prevalence of lateral incisor agenesis in the Maltese population is 3.21% (Camilleri and Mulligan, 2007). The high figure found in association with ectopic canines suggests that both traits are related, in keeping with previous evidence (Becker *et al.*, 1981; Svinhufvud *et al.*, 1988; Peck *et al.*, 1994; Pirinen *et al.*, 1996; Baccetti, 1998).

Pedigree Analysis

The pattern of transmission was that of an autosomal-dominant trait with low penetrance. Expression was highly variable both within and between families. Incomplete penetrance was evident, with obligate carriers unaffected by ectopic canines. This may be explained by the highly variable expressivity seen, with failure of penetrance in some cases.

Considering the high population prevalence, the figure for λ_R suggests a significant genetic component. When the drop-off in λ_R (decreasing degrees of relation) is plotted as a function of the Relative Risk, the gradual decrease for every decreasing degree of relationship is suggestive of a genetic contribution and does not rule out genetic heterogeneity between families. This analysis cannot distinguish between single-locus or multi-locus inheritance for ectopic canines (Risch, 1990). However, there is no difference in prevalence between families ascertained from one or two probands for first-, second-, or third-degree relatives. Nor is there a difference in the sib-sib/parent¬offspring relationships in the two types of families. This is indicative of a single gene, since multiple genes would be expected to give a higher prevalence in families ascertained from two probands.

There was a considerable sex bias toward females, and, surprisingly, elimination of the probands from the calculation produced an increase in the Female-to-Male ratio of the sample. There is no plausible explanation to date for this bias. However, there was no significant difference between the risks to relatives of male or female probands, as expected for an autosomal major gene. Were the trait to be multifactorial, the risk would be higher for the relatives of the less susceptible sex (Farrer and Cupples, 1998).

Table 2. Segregation Analysis Results: the -2 Log Likelihoods* of the Various Models Tested

Model	Allele Freq.	Penetrance AA	AB	BB	Heritability h2	-2 Log Likelihood	No. of Parameters	p-value	
Mixed dominant	0.89	0	0.32		0.47	379.76	4	-	0.09
Polygenic	-		0.12		0.4	384.49	2	0.07	
Sporadic	-		0.18		-	387.86	1		0.01
Dominant	0.89	0	0.36		-	380.34	3	1.00	
Co-dominant	0.89	0	0.36	0.36	-	380.34	4		0.09
Recessive	0.38	0		0.32	-	383.22	3	-	

* These were compared by the Likelihood Ratio test. P-value columns show pairwise tests of models in adjacent rows, with a significant p-value implying that the more complex model provides a better fit to the family data. The dominant model provides the best overall fit.

Segregation Analysis

The inheritance of ectopic canines in the families was modeled assuming a single major gene and with polygenic inheritance. The results of the segregation analysis implied that the sporadic model was rejected compared with all the genetic models. The dominant model fitted equally as well as the co-dominant

model, suggesting that a dominant model adequately describes inheritance in these pedigrees. The polygenic model was rejected, and there was no evidence for a polygenic component in addition to a dominantly inherited major gene. The low penetrance (36%) for the dominantly inherited gene allows for the existence of an additional environmental influence, which would determine which genetically susceptible individuals show the phenotype.

Twins

The inclusion of the first five sets of monozygotic and the first six sets of dizygotic twins was serendipitous, since selection was on the basis of affected probands only. The low concordance shown by the monozygotic twins is consistent with epigenetic or environmental factors influencing the eruption of teeth, and is at variance with the epidemiological evidence of a significant genetic component. Furthermore, the incidence of monozygotic twinning for the Maltese Islands is 4.5 *per* 1000 (Savona-Ventura and Grech, 1988). The prevalence of monozygotic twins selected consecutively in this sample was nearly double that number, and the difference was statistically significant ($p = 0.01$).

Monozygotic twins share identical DNA sequences; however, they are often discordant for certain phenotypes. Hypodontia, a genetic disorder associated with ectopic canines, is more prevalent in twins. These exhibit**(AQ)** a high degree of discordance (Keene, 1971; Lapter *et al.*, 1998; Townsend *et al.*, 2005). It is possible that epigenetic events may be responsible for discordant expression in genetically identical individuals. Indeed, certain imprinting-associated diseases, such as Beckwith-Wiedemann Syndrome, are associated with multiple births and discordance in monozygotic twins (Elliott and Maher, 1994). In contrast, ectopic movement of the maxillary canine has been detected in children as young as 5 years of age (McSherry and Richardson, 1999), and the rapidly growing child is susceptible to a multitude of factors that potentially affect intrabony movement of the canine. Evidence for environmental influence on epigenetic mechanisms further clouds the issue (Bird, 2007).

In conclusion, the evidence gathered from analysis of the pedigrees supports the hypothesis of a genetic etiology for ectopic canines, with a single locus being involved, although genetic heterogeneity across pedigrees cannot be ruled out in a modeling study. The most likely mode of transmission is autosomal-dominant. The incomplete penetrance of the dominant locus allows for an environmental component. Further investigation into the molecular etiology of ectopic canines is therefore justified, although the number of monozygotic twins in the sample, plus the discordance of these twins, raises the possibility of gene silencing.

ACKNOWLEDGMENTS

This research was funded jointly by the Faculty of Dental Surgery of the University of Malta (SC) and by the Dental Institute, King's College London (FM), with additional support from the Wellcome Trust (GR076024, to CML).

REFERENCES

Baccetti T (1998). A controlled study of associated dental anomalies. *Angle Orthod* 68:267-274.

Becker A, Smith P, Behar R (1981). The incidence of anomalous maxillary lateral incisors in relation to palatally-displaced cuspids. *Angle Orthod* 51:24-29.

Bird A (2007). Perceptions of epigenetics. *Nature* 447:396-398.

Bjerklin K, Kurol J, Valentin J (1992). Ectopic eruption of maxillary first permanent molars and association with other tooth and developmental disturbances. *Eur J Orthod* 14:369-375.

Blouet BW (2004). Malta and the Knights of St John. In: The story of Malta. Malta: Progress Press, pp. 81-122.

Camilleri S (1995). The prevalence of impacted permanent canines in Maltese schoolchildren—a pilot study. *Maltese Med J* 7:42-45.

Camilleri S, Mulligan K (2007). The prevalence of malocclusion in Maltese schoolchildren as measured by the Index of Orthodontic Treatment Need. *Malta Med J* 19:19-24.

Chaushu S, Sharabi S, Becker A (2002). Dental morphologic characteristics of normal versus delayed developing dentitions with palatally displaced canines. *Am J Orthod Dentofacial Orthop* 121:339-346.

Chaushu S, Zilberman Y, Becker A (2003). Maxillary incisor impaction and its relationship to canine displacement. *Am J Orthod Dentofacial Orthop* 124:144-150.

Davis PJ (1987). Hypodontia and hyperdontia of permanent teeth in Hong Kong schoolchildren. *Community Dent Oral Epidemiol* 15:218-220.

Elliott M, Maher ER (1994). Beckwith-Wiedemann syndrome. *J Med Genet* 31:560-564.

Ely NJ, Sherriff M, Cobourne MT (2006). Dental transposition as a disorder of genetic origin. *Eur J Orthod* 28:145-151.

Ericson S, Kurol J (1986). Radiographic assessment of maxillary canine eruption in children with clinical signs of eruption disturbance. *Eur J Orthod* 8:133-140.

Ericson S, Kurol J (1988). Early treatment of palatally erupting maxillary canines by extraction of the primary canines. *Eur J Orthod* 10:283-295.

Farrer L, Cupples L (1998). Determining the genetic component of a disease. In: Approaches to gene mapping in complex diseases. Haines JL, Pericak-Vance Margaret A, editors. New York, Chichester: Wiley Liss, Inc., pp. 93-129.

Hasstedt SJ (2005). jPAP: document-driven software for genetic analysis. *Genet Epidemiol* 29:255.

Keene HJ (1971). Epidemiologic study of tooth size variability in caries free naval recruits. *J Dent Res* 50:1331-1345.

Kotsomitis N, Freer TJ (1997). Inherited dental anomalies and abnormalities. *ASDC J Dent Child* 64:405-408.

Lapter M, Slai M, Skrinjaric I, Muretic Z (1998). Inheritance of hypodontia in twins. *Coll Antropol* 22:291-298.

McSherry P, Richardson A (1999). Ectopic eruption of the maxillary canine quantified in three dimensions on cephalometric radiographs between the ages of 5 and 15 years. *Eur J Orthod* 21:41-48.

Paschos E, Huth KC, Fassler H, Rudzki-Janson I (2005). Investigation of maxillary tooth sizes in patients with palatal canine displacement. *J Orofac Orthop* 66:288-298.

Peck S, Peck L, Kataja M (1994). The palatally displaced canine as a dental anomaly of genetic origin. *Angle Orthod* 64:249-256.

Pirinen S, Arte S, Apajalahti S (1996). Palatal displacement of canine is genetic and related to congenital absence of teeth. *J Dent Res* 75:1742-1746.

Power SM, Short MB (1993). An investigation into the response of palatally displaced canines to the removal of deciduous canines and an assessment of factors contributing to favourable eruption. *Br J Orthod* 20:215-223.

Risch N (1990). Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet* 46:222-228.

Rose JS (1966). A survey of congenitally missing teeth, excluding third molars, in 6000 orthodontic patients. *Dent Pract Dent Rec* 17:107-114.

Savona-Ventura C, Grech ES (1988). Multiple pregnancy in the Maltese population. *Int J Gynaecol Obstet* 26:41-50.

Svinhufvud E, Myllarniemi S, Norio R (1988). Dominant inheritance of tooth malpositions and their association to hypodontia. *Clin Genet* 34:373-381.

Townsend GC, Richards L, Hughes T, Pinkerton S, Schwerdt W (2005). Epigenetic influences may explain dental differences in monozygotic twin pairs. *Aust Dent J* 50:95-100.

Yilmaz HH, Turkkahraman H, Sayin MO (2005). Prevalence of tooth transpositions and associated dental anomalies in a Turkish population. *Dentomaxillofac Radiol* 34:32-35.